

From *confetti* to *kawaii*:

A study of the nativization of Italian and Japanese loanwords in The Oxford
English Dictionary

Ingeborg Vangsnes



Master's Thesis in English Linguistics

Department of Foreign Languages

University of Bergen

May 2019

Summary in Norwegian

Gjennom ein kultur-kontekstuell og språkleg analyse av lånord i The Oxford English Dictionary (OED) har denne oppgåva undersøkt trendar i integreringa av lånord frå italiensk og japansk i engelsk.

Oppgåva har gjort greie for den sosiopolitiske og kulturelle konteksten for språkleg kontakt mellom dei aktuelle kulturane og den engelskspråklege verda som dannar bakgrunnen for språkleg kontakt. Gjennom analysen konkluderer oppgåva med at forskjellar ved kontaktsituasjonen mellom dei to kjeldespråka ikkje har tydelege innverknader på typen endringar som er aktive i den språklege integreringsprosessen. Det som viser att frå dei to svært ulike kontaktsituasjonane, er ein forskjell i omfanget av lån, både gjennom talet på ord og ordklassane dei tilhøyrrer. Det vert også gjort greie for dei typologiske forskjellane mellom kjeldespråka italiensk og japansk, og dei moglege resultatane desse ulikeheitene mellom engelsk og kjeldespråka kan føre til i integreringa av lånord. Den språklege analysen har sett på trendar i fonologiske, ortografiske og morfologiske tilpassingar ved totalt 1832 lånord frå dei to kjeldespråka: 1306 frå italiensk og 530 frå japansk. Trendane viser at prosessane bak tilpassingar i integreringsprosessen i liten grad er sensitive til forskjellar i typologien til kjeldespråka, men i første rekke baserer seg på reglane i det engelske språksystemet.

I oppgåva vart det også prøvd ut ei hypotese om at ord som beheldt språklege trekk frå kjeldespråket, ville vere ord som var sjeldan brukt i engelsk, slik bruk er dokumentert gjennom såkalla «frequency bands» i OED. Det vart ikkje funne eit slikt tydeleg samsvar. Vidare viste oppgåva at japanske lånord fylgjer same mønster som italienske med trykk på nest siste staving. Analysen inkluderar også eit diakront perspektiv, ved å dele lånorda inn i to tidsperiodar: 1500 til 1700 og 1701 til i dag. Det vart ikkje funne tydelege forskjellar generelt i dei språklege tilpassingsprosessane eller nivået av språkleg integrering i samanlikninga av lånord frå dei to periodane.

Dei ulike innfallsvinklane og aspekta av analysen gir samla sett eit heilskapleg bilete av dei ulike faktorane som kan påverke importen av lånord og dei språklege prosessane som er involverte i integrasjonen av desse i engelsk. Samtidig viser oppgåva at det med eit breiare kjeldegrunnlag og andre metodar kan vere mogleg å få fram ny kunnskap på dette feltet.

Acknowledgements

There are many people deserving of thanks for helping me complete this project. First, I want to thank my supervisor professor Jerzy Nykiel for his patience and guidance throughout the process.

I also want to express my sincere gratitude to my flatmates, friends and fellow students for emotional as well as academic support and encouragement. The countless long nights of writing would not have been possible without your moral support and companionship. Special thanks are in order to the esteemed members of the writing group *heihei* for much appreciated feedback along the way.

Finally, this project could not have been completed without the continued support, encouragement and vital input on structure provided by my wonderful parents.

Table of Contents

1	Introduction	1
1.1	Aim and scope	1
1.1.1	Research questions and hypotheses	2
1.2	Structure of the thesis	4
2	Theory	6
2.1	Defining the loan word	6
2.2	Loan phonology: Two models of loaning	9
2.3	Orthographic effects	11
2.4	Language contact	12
2.5	Core-periphery organisation of the lexicon	14
2.6	Previous studies on borrowing and loanword adaptation	15
3	Methodology and data collection	19
3.1	Methodology: lexicological approach	19
3.2	Data collection and organisation.....	21
3.2.1	The limitations of the OED as a data source	24
3.3	Frequency.....	25
3.4	Factors for Analysis	27
4	Cultural analysis: history of contact	30
4.1	Contact between Italy and the English-speaking world.....	31
4.1.1	Early contact: 1500 – 1700.....	31
4.1.1.1	The Renaissance	32
4.1.1.2	Literary interest	34
4.1.1.3	Travel	35
4.1.1.4	Religion and politics.....	35

4.1.2	Later contact: 1700 onwards.....	36
4.1.2.1	Travel: The Grand Tour to modern tourism.....	36
4.1.2.2	Music, Art and Science	38
4.1.2.3	Italian cuisine and emigration to the US	39
4.2	Contact between Japan and the English-speaking world.....	40
4.2.1	Early contact.....	41
4.2.1.1	Trade.....	42
4.2.1.2	Japanese Isolationism (1639 – 1853)	43
4.2.2	Later contact	45
4.2.2.1	World wars and American Military Bases in Japan.....	46
4.2.2.2	Japanese cuisine	48
4.2.2.3	Business and technology	48
4.2.2.4	Japanese Popular Culture: “Cool Japan”	49
4.3	Preliminary conclusions.....	50
5	Linguistic features.....	53
5.1	Italian features.....	53
5.1.1	Phonemes of Italian	53
5.1.1.1	Vowels.....	53
5.1.1.2	Consonants	54
5.1.1.3	Notable variation	56
5.1.2	Syllable structure and stress pattern	56
5.1.3	Morphology: Inflection	57
5.1.4	Orthography.....	58
5.2	Japanese features.....	58
5.2.1	Phonemes of Japanese	59
5.2.1.1	Vowels.....	59

5.2.1.2	Consonants	60
5.2.1.3	Notable variation	61
5.2.2	Syllables and stress: mora and pitch accent.....	63
5.2.3	Morphology: inflection.....	65
5.2.4	Orthography.....	65
6	Linguistic analysis	67
6.1	Early loans: 1500 – 1700	67
6.1.1	Frequency	68
6.1.2	Phonological adaptations.....	70
6.1.3	Orthographic adaptations and orthographic effects.....	73
6.1.4	Morphological adaptations	76
6.1.5	Syllabic structure and stress pattern	79
6.2	Later loans: 1701 – the present	81
6.2.1	Frequency	83
6.2.2	Phonological adaptations.....	85
6.2.3	Orthographic adaptations.....	87
6.2.4	Morphological adaptations	88
6.2.5	Syllable distribution and stress pattern.....	91
6.3	Discussion of linguistic analysis.....	93
7	Conclusion.....	97
7.1	Avenues for further study	100
	References	102

Index of tables

Table 5.1: Italian Vowels.....	53
Table 5.2: Italian Consonants	55
Table 5.3: Japanese Vowels.....	59
Table 5.4: Japanese Consonants	60
Table 6.1: frequency distribution of early loans.....	68
Table 6.2: frequency distribution of early loans: percentages.....	69
Table 6.3: stress placement on early Italian loans	79
Table 6.4: Frequency distribution of late loans	84
Table 6.5: Frequency distribution of late loans: percentages	84
Table 6.6: Stress placement on late Italian loans.....	91
Table 6.7: Stress placement on late Japanese loans.....	93

1 Introduction

1.1 Aim and scope

The aim of the present thesis is to find trends in the nativization of Italian and Japanese loanwords attested in the Oxford English Dictionary (henceforth OED), considering factors such as frequency, typological similarities and sociocultural context in the adaptation of spelling, pronunciation, stress assignment and inflection. The thesis will include a comparative aspect, where the similarities and differences between trends in linguistic nativization between the two source languages will be discussed, in order to ascertain to what degree trends in adaptation can be said to be universal strategies, or sensitive to specific factors pertaining to the unique borrowing situations. The trends in linguistic nativization observed will be further discussed in terms of the linguistic contact situation and related to the OED frequency bands and the core-periphery theory of the lexicon. This theory considers the lexicon as consisting of different strata, where native core vocabulary follows a strict set of linguistic rules that more peripheral items, such as loanwords that remain markedly ‘foreign’, are ‘allowed’ to violate. A correlation is assumed between the core-periphery theory and a cline of high to low frequency of use, which the OED bands serve to indicate.

While several studies done on loanwords, and books on the history of English, have devoted time to the integration of words of Scandinavian, French and Latin origin, there have been few comparative studies done on the nativization of loanwords into English from other languages. The comprehensive look at borrowing into English by Durkin (2014) only devotes a handful of pages to borrowing from Italian and Japanese, while whole sections cover Scandinavian and Latin for instance. One central work on Italian loanwords is that of Pinnavaia (2001), who performed a lexicographical, linguistic and cultural analysis of the Italian borrowings in the second edition of the OED. Fournier (2018) conducted a study of the stress assignment on Italian loanwords in English compared to their native stress patterns in Italian. Based on his findings and those of other French linguists, he hypothesised that penultimate stress can be considered a ‘standard’ for foreign stress in English, a phenomenon he calls the “Italian rule”. While studies have been conducted on the stress patterns of French loanwords,

such as Svensson (2004), little appears to have been done on other languages. I have also not been able to find studies similar to Pinnavaia (2001) on Japanese loanwords, nor studies of loanwords in English that compare the nativization of loanwords from typologically different source languages. The present thesis thus aims to shed some light on these areas of loanword studies. I will test the Italian Stress Rule on the data from Japanese, to see if eligible constructions follow this rule in stress assignment. The comparative aspect of my thesis aims to add a new dimension to the study of the nativization of loanwords and allows for the discovery of cross-linguistic patterns in adaptation, in relation to the aforementioned variables of typological difference between the source language and the borrowing language as well as the historical contact situation.

1.1.1 Research questions and hypotheses

As outlined above, what my thesis aims to uncover are trends in the nativization of loanwords in English in terms of linguistic adaptation affecting phonology, morphology and orthography. The second question my thesis will attempt to supply an answer to is whether loanwords from Italian undergo different processes of adaptation than Japanese loanwords, based on typological considerations or as a result of factors in the linguistic contact situation. Another aspect concerns the basis for my approach to the theory of a core-periphery organisation of the lexicon, and the usefulness of textual frequency as represented by the OED frequency bands in the positioning of loanwords in this model of stratification of the lexicon from core vocabulary to a peripheral vocabulary. A concrete question that I aim to answer is related to stress assignment of Japanese loanwords and whether these will follow the stress rule postulated by earlier linguists or not. Finally, I want to include a diachronic perspective, and aim to discuss the role time of admission can be said to have played in nativization through comparing loans from the early period (1500 – 1700), which have stayed in the language for longer, with later loans (1701 – present). Based on the above, I have formulated the following six research questions (RQ):

RQ 1: Are there noticeable trends in the nativization of linguistic features in loanwords from Italian and Japanese in English?

RQ 2: Do differences in the cultural contact situation with the two source languages translate into differences in the linguistic adaptation of material in a borrowing situation?

RQ 3: Are the adaptation processes active in nativization sensitive to differences in source language typology?

RQ 4: Can the OED frequency bands be employed as a useful measure of the relative positioning of lexical items along a core-periphery organisation of the lexicon?

RQ 5: Will Japanese loans follow “the Italian rule” of penultimate stress?

RQ 6: Are there notable differences in the nativization of early (1500 – 1700) compared to later (1701-) loans that can point to time of admission as a factor in nativization?

The following hypotheses (H) were formed on the basis of the above research questions:

H 1: I expect to find certain clear trends in adaptation of linguistic features through nativization that base themselves on rules of the L1 phonology and orthographic convention.

H 2: There will be a noticeable connection between any differences in linguistic nativization processes and the history of cultural contact.

H 3: Regular processes of phonological and orthographic adaptation will apply to loanwords from both Italian and Japanese, as the L1 system will be more important in the nativization than the source language typology.

H 4: The OED frequency bands can be used as a measure of placement in a core to periphery hierarchy, as there will be strong correlations between the nativization of linguistic features and the frequency band.

H 5: A clear majority of Japanese loanwords will be stressed on the penultimate syllable, in line with the “Italian rule”.

H 6: Early loanwords will show a higher degree of nativization, and exhibit fewer markedly foreign features than later loanwords.

1.2 Structure of the thesis

Chapter 2 outlines the theoretical foundations for the thesis, including a discussion of the definition of the concept of ‘borrowing’, the principles of linguistic transmission within loan phonology, orthographic effects in loanword transmission and reproduction, language contact and a summary of the core-periphery theory in stratification of the lexicon by Itô and Mester (1995) as used by Friesner (2009). Previous studies on loanwords and their adaptation in English are also summarised.

The methodology and process of data collection for the present thesis are presented in chapter 3, including the process of narrowing down the data set, and the specific factors that will be used in the analysis will be outlined. The data has been divided into two sections based on the year of first attestation listed in the OED. Loanwords attested between 1500 and 1700 have been classified as ‘early loans’, while loanwords first attested after 1701 have been labelled ‘later loans’. This division was made to facilitate a diachronic perspective of analysis, to see if any differences can be discerned in the nativization processes that were active in the early period versus in later centuries.

The next section, chapter 4, provides a contextual analysis of the cultural contact facilitating linguistic transmission between the source cultures and the English-speaking world. The contact situations will be divided into an early and a later period, and these more or less correspond to the division of early versus later loans in my data¹. The cultural analysis of the history of contact serves the purpose of facilitating a comprehensive comparison, where cultural factors may help illuminate potential variation in the linguistic processes of the borrowing situation from Italian and from Japanese.

The linguistic features of the source languages are outlined in chapter 5, in order to explore the typological differences which might affect the reproduction of the loanwords from these languages in English and subsequent nativization within the English system. In addition to facilitating comparison between loanwords and their original etyma, this will help make any foreign features retained in loanwords obvious

¹ The division between early/late is not as clear in Japanese, for reasons which will be explained in chapter 4.

and allow for identification of words that have resisted nativization processes and remained as peripheral items of the lexicon.

The previous chapters lay the foundations for the analysis and discussion of my data in chapter 6. The analysis is divided into early and late loans, and covers the following categories: frequency, phonological adaptations, orthographic adaptations, morphological adaptations and syllabic structure and stress pattern. The chapter will conclude with a comparative section where the similarities and differences observed in the early versus later loans will be discussed.

2 Theory

This chapter outlines the main theoretical approaches I have used in my thesis. The first section discusses some definitions of the term *loanword*, and some of the problematic aspects with the definition I make use of in this paper. The following sections outline several different theoretical frameworks and approaches of linguistic study of loan words, including loan phonology, core-periphery theory of lexical strata, orthographic effect and language contact. Finally, I summarize some of the studies that I have based my project on and from which I have drawn important aspects of my approach.

2.1 Defining the loan word

In order to discuss and analyse linguistic borrowing, it is first useful to define the terms used and discuss some of the differences of opinion on the subject found in different disciplines of linguistics. The present thesis uses the terms *a loanword*, *a loan* and *a borrowing* to refer to the tokens themselves, and *borrowing*, *transmission* and occasionally *loaning* for the event. The variation is employed for the sake of avoiding repetition. For the process of linguistic integration of these words into English, the term *nativization* is used, in line with the terminology used in the field of loan phonology (by Calabrese and Wetzels (2009) among others) and the specific processes involved are referred to as *adaptations*. Other linguists, as well as the OED, have used terms such as *anglicization* or *naturalization* for the same phenomena, but these have not been used in the present thesis for the sake of consistency.

In contact linguistics, theorists such as Matras (2009) take issue with the basic semantics of the terms *loan* and *borrowing*. Matras (2009) prefers the term *replication*, and others in his field suggest *copying*, as terms that do not appear to imply ownership of words or suggest that elements taken from one language and used in another are going to be returned, an implication arguably present in *loan* and *borrowing* arguably do. However, as Haugen (1950) and others argue, as the most established terms, *borrowing* and *loan* are useful and will suffice for a discussion of these processes, despite the specifics of the semantic implications. Haugen (1950) specifies that while *borrowing* can be used generally, the term *loanword* is “ordinarily limited to terms ... in

which speakers have imported not only the meaning of the form but also its phonetic shape" (1950: 213). He thus separates borrowing into two types: *importation*, which involves imitation resulting in a form that would still be recognisable to a speaker of the source language, and *substitution*, where the form is no longer recognisable to a speaker of the source language, but has been substituted for a pattern in the borrowing language (Haugen 1950: 212). The latter would involve processes such as loan translation, where semantic content is borrowed but the form is substituted for a native one. The fact that a loanword is an imitation of a foreign model is a point that Gusmani (1973)² also stresses, and the present thesis will primarily concern itself with the type of borrowings that involve such reproductions of a recognizable phonetic shape.

However, deciding that the terms *loanword* and *borrowing* are suitable does not solve the question of how to define these terms, which is still a topic for much debate among linguists. When is a word a borrowing? One simple definition is given by loan phonologists Calabrese and Wetzels (2009), who state that "once [a] learned word [from another language] is uttered publicly or even silently by the speaker to himself, it is a loanword" (2009: 2). In broad terms, their definition states that once a word from a foreign language is uttered by a single person in the context of that speaker's native language (L1), it can be considered a loanword. Haugen (1950) gives a similar definition, claiming that "the attempted reproduction in one language of patterns previously found in another" (1950: 212) is what constitutes a borrowing.

Other linguists argue that a certain spread and frequency of use among more speakers is necessary for a word to be properly considered a loan, but Matras (2009) points out that no uniform standard has been established according to which the frequency of occurrence can be judged. The domains in which these words are used are also factors. How can one usefully and clearly differentiate between foreign words being used referentially or to achieve a certain effect and loans that are fully integrated into L1? One distinction to make here is that between bilingual codeswitching and borrowing. In general terms, borrowing is used to describe the "diachronic process by which languages enhance their vocabulary" (Matras 2009: 106), while the term codeswitching is used for "instances of spontaneous insertions in the speech of

² In translation to English by Pinnavaia (2001)

bilinguals” (2009: 106). Linguists still disagree on where to draw a sharp divide between these two concepts, and questions of how long a word must be in use to be considered a borrowing is hard to define, but Matras (2009) instead presents the issue as a continuum, with codeswitching on one end, and borrowing on the other. Among the dimensions of this continuum are *regularity*: ranging from single occurrences to regular use, *structural integration*: ranging from not linguistically integrated to integrated into the L1, *functionality*: from special conversational effects or stylistic choice to default expression, and importantly the *bilinguality* of the speakers, ranging from bilingual to monolingual (Matras 2009: 111). According to Matras, these dimensions more holistically represent the various constraints and preferences that condition the employment of a lexical item in a variety of contexts, and that contribute to separating codeswitching from borrowing. The *bilinguality* dimension is of particular importance to the model, as codeswitching first and foremost involves bilinguals with a high degree of fluency in both the source and recipient language. Once a word has been established as a borrowing on the other hand, its usage does not presuppose such advanced levels of knowledge of the source language from which a word was originally taken. Matras (2009) assumes that bilinguals are also essential in the initial stages of borrowing, and so in such early stages, the two concepts can be difficult to distinguish. The division is clearly a complex one, but one aspect that can help in forming a practical distinction between the concepts is the domains of usage. The concept of bilingual codeswitching is often mainly discussed as a conversational phenomenon, where usage of insertions and codeswitches is sensitive to the levels of bilingualism of the involved interlocutors. One could thus tentatively conclude that items recorded in writing are more likely to represent instances of borrowing than one-word codeswitches. Matras (2009) concludes that “borrowing involves a long-term or permanent licence to lift selection constraints on the use of a word-form or structure” (2009: 147), meaning that a form that was restricted to specific conversational domains involving bilinguals (codeswitching) has been extended to wider usage in which monolinguals are also involved.

For the purposes of the current project, the practical distinction between foreign word and loan word is defined by the data source: a loan word is a word originating in a different language than English, as listed by the OED. As my approach is based on pronunciation, my data will not include what Haugen (1950) called *substitutions*, such

as calques, i.e. loan translations (such as *flea market* from French *marché aux puces*), but only words whose form still reflects the form of the source language etymon.

As seen in the discussion above, it would appear that Matras' (2009) theories and approaches assume that borrowing is a phenomenon linked closely to bilingualism, and that the initial act of borrowing is carried out by bilinguals. Within the field of loan phonology however, many linguists would strongly disagree with that as a premise, as will be discussed in the following section.

2.2 Loan phonology: Two models of loaning

A central field for a linguistic analysis of loanwords is that of loan phonology. Through studying the nativization of loanwords phonologists can gain insight into phonological processes of language generally and have thus shown keen interest in the processes involved in borrowing. Calabrese and Wetzels (2009) outline how in the field of loanword phonology, different linguists tend to consider the process of loaning in terms of two different scenarios, which roughly correspond to two current theoretical models of loanword phonology. These are commonly referred to as the *perception model* (Calabrese and Wetzels 2009), which involves a scenario of borrowing with nativization through perception, i.e. where adaptations to fit L1 systems happens during input, and the *phonetic stance model* which corresponds to a process of nativization through production, i.e. where adaptations occur during output.

In the nativization-through-perception scenario, borrowing is implemented by a speaker that fills a gap in his language by taking a word from another language which they know poorly or not at all. In this case they first need to learn the relevant word. The model assumes that once the learned word is uttered publicly or silently by the speaker to themselves, it can be considered a loanword. Given that the speaker does not speak the second language well, the word will display adjustments and adaptations in order to conform to the speakers' L1 output rules (Optimality Theory: constraints). The hypothesis is that these modifications have already occurred during perception and learning of the word from the source language. The initial input in this scenario is the acoustic signal that the speaker hears when first exposed to the word, which is then filtered through the phonological rules and constraints of their L1.

In the nativization-through-production scenario, borrowing is thought to occur when a bilingual speaker fills a gap in one of the languages that they know by taking a word from the other language they are fluent in. The model thus assumes that the speaker retrieves the underlying representation of the borrowed word from their mental dictionary for L2 and generates its surface representation while speaking in L1. If the surface representation of the word is generated by using the phonological, or more generally the grammatical system of L1, the word undergoes adaptations and adjustments and is nativized according to the grammar of L1. Another alternative is that the surface representation of the word is generated by using the L2 grammatical system. In this case, the word would be reproduced and pronounced in its proper L2 shape.

To sum up the stances presented by the papers in the collection of studies on loan phonology by Calabrese and Wetzels (2009), the majority conclude that the nativization of loanwords occurs under the nativization-through-perception scenario, which supports the perception model. It might then be tempting to assume all nativization occurs during perception, but this would be implausible cross-linguistically. Bilinguals of varying degrees are clearly also a force in many contact situations, and to completely dismiss their underlying knowledge of L1 and L2 structure would be reductive. However, Calabrese and Wetzels (2009) conclude that the models can be unified. They suggest that the input to nativization is always phonetic, i.e. the word as heard – the acoustic signal, while the treatment is always phonological and can occur either during perception or during production. They maintain that perception and production cannot be separated in the study of nativization of loan words.

As the name *loan phonology* suggests, these theories all operate with the presumption that borrowing and subsequent adaptation can be explained in phonology, as it is fundamentally a phonological process, based on the input of acoustic information. However, linguists in other fields propose that other aspects are also involved. Friesner (2009) suggests that in order to get a full picture of nativization, one must look not only at different modules like phonology and morphology, but at linguistic differences on a typological level, as well as external explanations such as orthography and socio-political and cultural factors, which also have an impact on nativization patterns.

2.3 Orthographic effects

The term *orthographic effects* refers to the influence or interference effects spelling can have on the adaptation of loanwords, specifically on the adaptation of pronunciation. Vendelin and Peperkamp (2006) assert that many factors besides the purely grammatical are likely to influence the adaptation of loanwords and point to orthography as one such factor. According to them, orthography is usually described as marginal, if taken into account at all, in studies on loanword adaptation. They conducted a study with French-English bilingual speakers to find out how different input would impact the reproduction of vowels in loanwords from English. In this study Vendelin and Peperkamp (2006) establish a distinction between two types of orthographic effect; ‘reading’ adaptations and adaptations based on between-language grapheme-to-phoneme correspondence rules (2006: 3) The first type refers to instances where speakers read a loanword as if it was a native word in their L1 i.e. according to conventions for pronunciation of graphemes in the L1. The second type assumes a level of standardisation as to the way source language graphemes are pronounced in the borrowing language, such as how French speakers are taught in school to pronounce the English grapheme <oo> as their native vowel /u/ (Vendelin and Peperkamp 2006: 997).

Their study of English loanwords in French shows that in borrowing situations with mixed (i.e. oral + written) input, French speakers more often produce output which reflects the way they are used to reading English graphemes. Vendelin and Peperkamp (2006) thus conclude that loanword adaptations are sensitive to the presence versus absence of a written representation of the word. They also maintain that an experimental framework is best suited for studies of loanword adaptation, as orthographic effect is difficult to determine, and even in an experimental setting they found it hard to quantify the exact influence of orthography on pronunciation. This was in part due to the fact that it in many cases can be hard to distinguish adaptations based on orthography from adaptations based on underlying phonological and/or phonetic similarity. Recreating an experimental setting including respondents exposed to various input following the methods of Vendelin and Peperkamp (2006) is however beyond the scope of the present study. I will still attempt to make arguments for cases where orthographic effects are likely in my dictionary-based approach.

Coetsem (1988) also underlines how both acoustic and visual aspects of language can be involved in language contact and interact intimately. Coetsem (1988) claims that speakers are motivated by a tendency towards congruity between spelling and pronunciation, which can lead to two distinct processes in a nativization situation. The first process is spelling-induced pronunciation, which refers to what others have termed *spelling pronunciation*, and also corresponds to what Vendelin and Peperkamp (2006) called *reading adaptations*. Coetsem (1988) uses the pronunciation of the English loanword *score* in Dutch as an example, as Dutch speakers tend to pronounce word-final <e> in their native system according to their spelling rules: [sko:rə] (Coetsem 1988: 93). The other process is pronunciation-induced spelling, whereby the source language pronunciation may first be adapted to that of the borrowing language, and the spelling of the loanwords is subsequently adjusted to better reflect the adapted pronunciation. This can be illustrated by a hypothetical scenario in which the Dutch speakers from the previous example had dropped the final <e> in order to reflect a pronunciation closer to the source language etymon.

In relation to my own data, I expect orthographic effects to be particularly relevant to Japanese loanwords, as the contact situation being one of limited direct contact (as will be discussed in chapter 4) means there is less likelihood of continued direct contact with acoustic signal. The main reference point will thus rather be the written representation of these words as adapted by the first borrowers. Italian words may also display some of these effects but based on geographic proximity and other factors of the contact situation, the increased exposure to the acoustic signal, might reduce the orthographic effects here.

2.4 Language contact

Another central field of linguistic study relevant to the analysis of borrowing is that of language contact. Matras (2009) states that in addition to linguistic factors, one must assume that borrowing also depends on extralinguistic factors such as the duration and intensity of cultural contact, the roles and status of the languages involved, as well as language attitudes. The intensity of language contact thus impacts the level of linguistic transmission and the type of linguistic content being transferred from one language to

another. Coetsem (1988) presents the borrowability of certain categories in a hierarchy related to the *stability* of language components or domains. According to such a hierarchy, elements involving a more stable domain, such as phonology, are more resistant to contact-induced change than a less stable domain such as vocabulary (Coetsem 1988: 3). Speakers are thus likely to preserve their more stable domains such as their L1 phonology when importing elements into the less stable domain of vocabulary.

Several more detailed hierarchies have been proposed for the types of linguistic content that can be transferred in a contact situation depending on the levels and intensity of cultural contact. Thomason and Kaufman (1992) established a borrowing scale, with contact ranging from casual to intense contact, and linguistic content types sorted into 5 categories along the scale. On the casual contact end of the scale they placed the lower category, which included content words, followed by function words and minor phonological and lexical semantic features in category 2. Features such as phonemes, word order and significant typological disruption were placed in the higher categories, which required an intense level of contact to be transferred. Matras (2009) problematises this model and points out the limitations of applying it to concrete language contact situations as the “intensity of contacts and cultural pressure are not strictly linear” (2009: 156). The types of contact or “cultural pressure” will also presumably differ across fields of discourse, as the languages involved in contact can have different roles in various domains of interaction. The levels of bilingualism, institutional support and community attitudes to the languages are also important factors that can influence borrowing. Matras (2009) further outlines how different case studies have confirmed a greater likelihood for open-class items (i.e. content words such as nouns) to be borrowed as opposed to closed-class items (i.e. function words). A crosslinguistic comparative study of languages in contact found that the categories most often involved in borrowing were nouns and conjunctions, followed by verbs, discourse markers and adjectives. Pronouns and inflectional affixes were found at the end of the hierarchy as the least frequently borrowed categories (Matras 2007).

A further key distinction in the definition of loanwords and the loaning process in language contact studies concerns the motivation behind the transmission of content from one language to another. Coetsem (1988) establishes a distinction between

borrowing and *imposition*. He defines borrowing as ‘recipient language agentivity’, meaning the act of transmission is due to L1 speakers copying a form from an L2 and consciously importing it into their L1. Imposition on the other hand involves ‘source language agentivity’, in which an L1 speaker reproduces elements from their L1 in the production of L2.

The motivations for borrowing are also relevant to consider and might affect the rate and degree of nativization. Matras (2009) presents the two most frequently cited motivations for borrowing as *gaps* in the inventory of the borrowing language and the *prestige* of the donor language. Typical examples of “gap-fillers” are words that accompany new concepts, such as new social activities, cultural acquisitions, community institutions or new innovations. These are often referred to as *cultural loans*. Rather than denoting new referents however, the typical *prestige*-loans on the other hand often have corresponding expressions in the borrowing language. Their use is thus not motivated by an aim to fill a gap of lexical reference, but rather serve the purpose of evoking special conversational effects and associations. The *prestige*-type of borrowing reflects the desire to imitate elements from communities that are socially more powerful, dominant or influential in order to gain social status through association. In time, such borrowings can also replace native expressions, as seen in the case of the word *uncle*, originally derived from French, which eventually replaced the Anglo-Saxon word *eam* (Matras 2009: 150).

2.5 Core-periphery organisation of the lexicon

A central study from which the present thesis borrows ideas and theoretic framework is Friesner (2009), who examines different factors affecting loanword adaptation through a set of loanwords from Turkish and French into Romanian. Friesner (2009) discusses the adaptation of these loanwords in terms of a 'core-periphery' organisation of the lexicon based on a model by Itô and Mester (1995). This model suggests that more peripheral lexical items may be "exceptional with regards to certain constraints of the source language" (Friesner 2009: 115) i.e. that they may retain features which violate certain constraints of the borrowing language which the core vocabulary items cannot violate. Friesner (2009) points out that the typical path for a foreign borrowing is to enter the

language in the periphery, as a marked foreign borrowing, and then optionally move towards full or gradual nativization “by changing its surface form to obey the previously violated constraints” (Friesner 2009: 116). Among the words in the peripheral sphere of the lexicon Friesner (2009) lists proper names, onomatopoeic forms and highly specialized vocabulary, in addition to foreign borrowings.

Friesner (2009) suggests that some languages might have more distinctly separated strata of core and periphery vocabulary, due to a history of contact with other languages and ‘waves’ of large-scale borrowing. Both Japanese and Korean are such languages, which in addition to a core native vocabulary have a large portion of ‘Sino-Japanese’ and ‘Sino-Korean’ vocabulary due to intense linguistic contact with Chinese historically. In more recent decades these languages have also loaned significant numbers of new words from other languages such as English, and these loans are still in various stages in the process of nativization, forming a third stratum. These words are thus considered more ‘peripheral’ than the vocabulary of Chinese origin, which is closer to the core and not perceived as markedly “foreign” in the same way anymore. The situation for English is somewhat more complex, as the waves of borrowing are not always so clearly distinguishable in terms of features.

Without going so far as to suggest a full model or structure for English, one could position certain waves in relation to each other however. Historically, Scandinavian loanwords could be considered slightly closer to the core than French and Latin loanwords given their time of integration, which again may be considered closer to the core than the more recent Italian loans and loanwords from other languages such as Japanese. More recent borrowings into English are likely to be considered more peripheral, but the degree to which they nativize and move closer to the core of the English lexicon depends on both internal linguistic factors such as how typologically “foreign” they appear and external factors such as cultural contact and presumably also the cultural associations of specific words, which influence how “foreign” they are perceived to be.

2.6 Previous studies on borrowing and loanword adaptation

An important previous study of loanwords and loanword adaptation is the book *The Italian Borrowings in the Oxford English Dictionary: A lexicographical, linguistic and*

cultural analysis by Pinnavaia (2001). The book represents a comprehensive study of several important factors of Italian borrowings listed in the OED, using methods of lexicographical, lexicological, and cultural analysis.

The lexicographical analysis includes an etymological analysis and a linguistic analysis. In her etymological study Pinnavaia (2001) explores the etymological metalanguage the OED uses to describe the relationship between the English lemmas and their Italian etymons and sorts these descriptions by the type of relationship they imply. She also devotes time to the limitations of this etymological metalanguage based on etymological errors found in the OED's data, and concludes that the inconsistencies in the descriptions of etymology and the lack of a key to explain etymological labels in the second edition of the OED comprises some of its main faults and shortcomings. Another aspect of loanword use and integration that Pinnavaia (2001) devotes time to is an analysis of the semantic functions of the Italian borrowings in the OED, and causes for obsolescence of certain words over time, as they relate to fields of discourse no longer as relevant in a modern society.

The linguistic analysis of Pinnavaia (2001) includes an overview of primary adaptations in the form of phonetic, orthographic, morphosyntactic, lexical and semantic adaptations, in addition to secondary adaptations of the same categories. She here finds certain noticeable trends in the phonological adaptations of Italian loanwords, particularly in weak syllables. Pinnavaia (2001) concludes her study with an account of the relations between semantic fields of loanwords and the historical background of cultural relations between Italy and England. Her cultural analysis is divided into periods spanning roughly two centuries each, stretching from as early as 1300-1500 to the latest period which covers the 1950s to the present. Pinnavaia's focus is on the interaction between the semantic fields of words and the areas of cultural contact these imply. She concludes that "to know the history of man's words is to reach the heart of human language and culture" (Pinnavaia 2001: 181). The intention of her study of the loanwords from Italian was to become better acquainted with the influence of Italian on English language and culture via the analysis of these Italian loanwords, and also to evaluate the OED's standing as a "lexical and etymological catalogue" of the evolution of the English language (2001: 181).

Another study central to the current thesis is that of Fournier (2018), which involved an analysis of stress assignment on Italian loanwords in English. Fournier (2018) compared the original stress patterns of Italian words with the stress placement of these words by English speakers after being borrowed into English. Fournier (2018) limited his data to words that maintained the form and segmental information of the etyma, and excluded cases where re-interpretation of syllable boundaries and the like could influence stress placement.

Fournier's study consisted of two main parts. The first section used a dictionary-based approach to analyse the differences in the stress patterns of Italian loanwords in English compared to their original etyma in Italian as recorded in the Oxford English Dictionary, the Longman Pronunciation Dictionary and the Cambridge English Pronunciation Dictionary. Fournier (2018) concluded that the stress pattern of the source language is "almost always" reproduced by the target language, as his dictionary-based study found an almost complete match between the stress pattern of the source language words and the corresponding loanwords. In addition to 9 partial matches, only 4 out of the total 309 tokens in his study did not display any stress match between Italian and English.

The second part of Fournier's study uses an experimental setting with seven English native speakers to test their sensitivity to Italian word stress. Fournier (2018) found that sensitivity to Italian stress was high, with most words reproduced with the correct stress pattern. However, there was a slight tendency to over-apply the penultimate stress pattern on words with other patterns in Italian, particularly among the speakers with no prior knowledge of Italian. Fournier thus proposed that further tests with more subjects with noticeably different linguistic backgrounds and proficiency in Italian could be interesting and illuminate the issue further. Differences in linguistic background of speakers could also be an interesting avenue for the study of the stress placement on loanwords in English.

As the study finds a strong tendency to stress Italian loanwords on the penultimate syllable, which reflects the reproduction of this stress pattern from Italian, the high number of loans with this pattern coming in from Italian has been proposed as a reason for penultimate stress placement on loanwords from other source languages in

English. Fournier (2018) points to the postulation of a pedagogical “Italian rule” by French phonologists, which states that words following the typical syllable structure of Italian loanwords, i.e. ending in an alveolar consonant followed by a pronounced vowel are stressed on the penultimate syllable.

3 Methodology and data collection

The following chapter outlines the data and process of data collection and methodology of the present thesis, as well as presenting the labels and categories used in the analysis. The advantages and limitations of the dictionary-based approach used by Pinnavaia (2001) and Fournier (2018) are outlined, as well as other challenges I have encountered in the data collection for the project. Finally, the background for the *frequency bands* of the OED and the typical characteristics of words in the different bands are explained.

3.1 Methodology: lexicological approach

The present study is a lexicological dictionary-based study of loanwords in English, and the data has been collected primarily from the Oxford English Dictionary online (OED). The online dictionaries Jisho.org (<https://jisho.org/>) and Grande Dizionario Hoepli Italiano (<https://dizionari.repubblica.it/Italiano>) have been used, for Japanese and Italian respectively, for comparisons between a loanword and their original etymon.

The term *lexicological analysis* is used by Pinnavaia (2001) to describe her method, and Hartmann and Stork (1972) define ‘lexicology’ as “a study and analysis of the vocabulary items of a language as well as their meanings and evolution” (1972: 129). My study is mainly synchronic in its approach, but as time of admission into the language as well as historical spelling variants are elements considered in the analysis, it also has a diachronic aspect.

As outlined in my theory chapter, the current thesis uses the theoretical framework of loan phonology (Calabrese and Wetzels 2009) as well as drawing on studies of orthographic effect (Vendelin and Peperkamp 2006) on loan adaptation, and the theory of a core-periphery organisation of the lexicon, as outlined by Friesner (2009) and Itô and Mester (1995). General theories of language contact have also been used to draw conclusions about the context for loanword adaptations in my data.

Building on the methods of the comprehensive study of Italian loanwords in the OED by Pinnavaia (2001), my study aims to find trends in the adaptation of loanwords from Italian and Japanese, as well as including a comparative aspect of the nativization

of words from the two languages. The methodological procedures Pinnavaia (2001) made use of in her work included both lexicographical and lexicological analysis. Her lexicographical analysis involved finding borrowings in the OED listed with Italian etyma. With the search functions available in the current online version of the OED however, this part of the analysis has not been necessary to replicate in full, but some manual selection has been done, as detailed in 3.2. Pinnavaia's lexicological analysis had three parts: the first being an etymological aspect, where she compared the etymological information given by the OED with that of other dictionaries. Secondly, she analysed linguistic factors such as pronunciation, spelling, morphosyntax and semantics in order to examine similarities and differences between the loan in English and the original Italian etymon. The third part of the lexicological analysis presented by Pinnavaia (2001) was a cultural analysis, which included detailing the background of cultural contact and subsequent linguistic contact between Italy and the UK historically.

The present study builds on and uses some of Pinnavaia's methodology, including some of her categories for the linguistic analysis, and the cultural analysis. Pinnavaia's cultural analysis has been an important inspiration and a central source for my cultural analysis of the history of language contact with Italian, and a model for my analysis of the contact with Japanese. However, my approach to the cultural analysis radically differs from that of Pinnavaia in one central aspect. While her focus was on how loanwords could illuminate areas of cultural contact, my study considers cultural contact as a background and basis for linguistic transmission. I further expect differences in the cultural contact to have consequences for the nativization processes that loanwords undergo. I have also included some sections on the more recent contact situation, including the contact between Italian and English in the United States, which Pinnavaia (2001) did not cover. I have not however included an in-depth etymological component in the current thesis, as this would be a very complicated and time-consuming process that would be beyond the scope of my project. My scope is thus limited to the etymological data provided by the OED alone, which arguably limits the project in certain ways. While the OED is widely held as the accepted authority on the English language, there may be missing or incomplete information for certain entries, as Pinnavaia (2001) concluded in her study. I acknowledge that based on the sources of my data, the conclusions I form will be similarly limited. It is not the purpose of this

thesis to make sweeping generalisations about loanwords per se based on this data, but only loanwords as recorded in the OED.

In addition to Pinnavaia (2001), I have drawn inspiration from Fournier (2018) and his study of stress assignment in Italian loanwords in English. He compared the stress pattern of Italian loanwords in English with the original stress pattern of the etyma in the source language. The present thesis similarly aims to find trends in the stressing of Italian loanwords, as well as comparing the pattern of stress assigned to Japanese loans. Additionally, the present thesis aims to test the “Italian rule” of penultimate stress that Fournier (2018: 13) states may have influenced other foreign loans with the same phonotactic structure.

3.2 Data collection and organisation

As discussed in 2.1, the definitions of what constitutes a loanword differ among linguists. The working definition used for the present thesis is one based on the simple definition given by Calabrese and Wetzels (2009): “once [a] learned word [from another language] is uttered publicly or even silently by the speaker to himself, it is a loanword” (2009: 2). Rather than when “uttered”, the present thesis considers a word a loanword when it is recorded in the OED as “borrowed from” or “based on” an etymon from a different language, and my data concerns borrowings from Italian and Japanese specifically. As a central factor in my analysis is adaptations to pronunciation, my data does not include loan translations. Some loan blends have been included, where the root of a word is based on a borrowed element, which has been combined with an affix productive in English. Some of these, as will be discussed further in chapter 6, are rather examples of re-interpretation of source language endings based on analogy with such productive affixes in the L1.

As discussed in 2.1, some linguists argue that a certain spread and frequency of use among speakers is necessary for a word to be considered a loan. The OED frequency bands are useful synchronic indicators of how widespread the use of a word is, as will be further discussed in 3.3. However, in terms of diachronic usage there are fewer options for effectively ascertaining which levels of usage and propagation

specific words had among the general English-speaking public and narrowing down the data based on such criteria has thus not been attempted in the current project.

The primary data for this thesis has been collected from the OED using the ‘advanced search’ option to apply filters in order to produce the results relevant for the scope of my project. The time period was narrowed down to 1500 – to the present day, and independent searches were conducted by language of origin for Japanese and Italian respectively. The results produced by these searches were manually copied over into a table in an excel spreadsheet. The results for loans of Italian origin had to be further narrowed down during the manual collection by excluding words with mixed or uncertain origins. I chose to not include words listed as “partly a borrowing from French, partly from Italian” or as “common Romance” and the like in order to ensure that their form and any adaptations that may have occurred are not in fact due to origin in other, similar linguistic systems, such as French, Latin or Spanish. Though this might have been an interesting in-depth etymological study of each entry, such an approach was not the focus of the present thesis. However, I have kept some early Japanese loans that were listed as being possibly borrowed through Portuguese. The Japanese data from the early period is scarce, and so the cases in question are few enough to keep track of. A few tokens from Japanese were left out as they had been formed in English through conversion rather than through borrowing³. Entries marked as ‘obsolete’ or ‘archaic’ were also left out, as these are no longer considered words in use, and thus no longer functional loanwords. Words tagged as ‘rare’ or ‘historical’ were included, as these still see some use, albeit limited. There were also a number of tokens listed without IPA transcriptions, and as pronunciation forms a central part of my analysis, these were also left out.

After the manual selection based on the criteria above have been carried out, my final data set consists of a total of 1832 tokens. For words with Italian origin, the OED advanced search option yields 1722 tokens, which I reduced to 1306 tokens. Of these, 307 were first attested in the period 1500 – 1700, while 999 tokens have first attestations after 1701. The OED lists 530 loanwords originating from Japanese, and

³ The noun *Nippon* ‘Japanese name for the country of Japan’ had several conversions, such as *Nipponese* (*adj, n*), *Nipponize* (*v*) and *Nipponized* (*adj*). Only *Nippon* actually involved the process of borrowing, however.

this number has been reduced to 526 in my data. 25 of these tokens are first attested before 1700, while 501 have first attestations after that.

The information taken from the OED includes information on year of first attestation in English, spelling variation, pronunciation, semantic definition, inflection, stress assignment and frequency. The ‘year of first attestation’ column of my table allows me to sort the words according to the time of admission into English, which is useful when trying to uncover any trends in nativization over time for a diachronic perspective on nativization. Information about spelling variation, including the relevant century or centuries in which these variants have been attested was included where such information was provided. The OED claims to list all variants they find, but that is of course no complete guarantee. Information regarding pronunciation, including any and all variant pronunciations listed were also copied from the OED into my corpus. Most entries are listed with a pronunciation using IPA for both British English and American English. According to their website “The Dictionary does not aim to cover dialectal variation in pronunciation within each variety”, and so my study will be limited to General American (GA) for the US variant and Received Pronunciation (RP) for the British variant and will not cover other English varieties. For some entries only one pronunciation was listed, with no distinction between GA and RP and no indication as to which one it was closest to. Inflectional information, which mostly constitutes plural marking for nouns, was also included in my table for entries where such details were specified, which was far from all. Historical spelling variants and the relevant centuries these were attested were also included where they were listed, which again was not all entries. Based on the pronunciation information supplied by the OED, I have tagged each word with the number of syllables it contains. However, a disclaimer is in order here, as this division is based on my intuition and knowledge of syllable restrictions in English, and it is possible that it may in certain cases be inaccurate. Any potentially relevant additional information was added in a ‘notes’ section in my table.

I also tagged each word with a broad semantic category based on the meaning supplied by the OED. The different broad semantic categories were based loosely on the categories of Serjeantson (1935), and are as follows: *Art*, (painting and sculptural art etc.) *architecture*, *culinary terms* (cooking techniques and food; dishes or ingredients etc.), *culture* (very broad, includes cultural practices specific to the country of the

source language and material culture), *music* (also including opera), *martial arts* (only relevant for Japanese; karate, judo; including titles and moves used), *military* (weapons, titles, ranks etc.) *miscellaneous* (category for cases that were hard to classify in broad terms), *nature* (names of plants, animals or physical geography), *geology and mineralogy*, *religious terms* (words related to religion or the practice thereof; shrines, titles of practitioners, festivals etc.), *science*, *society* (includes people, titles, corporations, institutions, etc.). Interjections, verbs, adverbs, adjectives and other part of speech categories have also been labelled for ease of retrieval later, and to separate them from the nouns that make up the main portion of the dataset.

The year of attestation as well as a short definition have been included when examples from my data have been used in the cultural analysis. In the linguistic analysis however, definitions and years have not been included unless specifically relevant to a specific point. Definitions have mostly not been deemed necessary for words above frequency band 4, as these are terms considered by the OED as known to most speakers.

3.2.1 *The limitations of the OED as a data source*

As mentioned earlier in this chapter, the information gathered from the OED varies from entry to entry, and as such it is relevant to point out that the OED has some limitations as a data source. One issue is that of first attestations. It is perhaps tempting to conflate a first attestation with the start of widespread usage of a loanword, but this is not the function of this information. The first attestation is just that – the first recorded time a word was attested in English text. The importance of keeping this in mind is shown perhaps most clearly in early Japanese loans, particularly those first attested in Kaempfer's *History of Japan* from 1727. As Durkin (2014) points out (see 3.2.2.2) many of these loans would only be used in a very limited context at such an early time and have not been recorded again in any kind of extended usage until centuries later. On the other hand, it is entirely possible that certain words were in use in areas of informal or oral discourse for a time before they were first attested in writing in the kind of sources consulted by the etymological researchers of the OED.

Another thing to keep in mind is that the OED is constantly being updated. Some entries have the information of its last update specified, and the years differ by a great deal. The second print edition of the OED was published in 1989, and some

entries have not been updated since that time. The OED states that the updating of the information in the online edition is ongoing, and that entries from M to R have been updated, as well as entries in “small but significant ranges elsewhere in the alphabet” (OED: frequently asked questions). There are thus differences between entries in the level of detail recorded, which might be down to the current work on updating the dictionary or reflect the fact that less etymological information is known about certain words. However, certain entries lack information that noticeably should have been there. One such example is frequency information, which is missing on a loanword like *sumo* from Japanese. This can hardly be because the word is too infrequent to have the frequency band information be calculable, and it thus appears that some entries lack information for reasons which are not related to their status or usage in the language. The hypothesis that the OED frequency bands can be used as a measure of the core-periphery organisation of the lexicon is thus met with some challenges, as the frequency bands are limited by other factors, as will be discussed in more detail in the following section.

3.3 Frequency

Information about textual frequency was gathered from the OED in the form of the ‘frequency band’ listed for each word. These frequency bands include eight tiers, ranging from 1 (very low-frequency) to 8 (very high-frequency), where each tier is about ten times more frequent than the previous. The bands are based primarily on textual frequency data from Google Books Ngrams data, which according to the OED was re-analysed in order to deal with ambiguities such as homographs. Spelling variation was dealt with by summing the frequency of each variant, and plural forms of nouns as well as verb tenses were similarly combined. The frequency represents the overall frequency of a word from 1970- to the present decade, summing the average for each decade into a total score. However, a notable limitation of the frequency bands is that the Google Books Ngrams data was generated in 2012. For the most recent loans, this means that the allocated frequency band may not accurately provide an up-to-date representation of frequency in current use. This might particularly effect recent loanwords concerning digital culture or recent cultural phenomena that may not have been in widespread public textual use before 2012, such as *emoji* from Japanese which is listed as band 1.

The lowest frequency (band 1) words constitute 18% of the total entries of the OED. These words are considered extremely rare, “unlikely ever to appear in modern text” and are restricted to obscure technical terms and terms with occasional historical usage. Some examples from my data include *gondoleet* ‘small boat or gondola’ and *guglio* ‘obelisk; needle’ from Italian and *joruri* ‘puppet theatre’ and *katsuramono* ‘category of Noh play’ from Japanese. Band 2 is by far the largest group and 45% of all words listed in the OED belong in this band. Band 2 words occur fewer than 0.01 times per million words, and the OED thus classifies them as “almost exclusively terms which are not part of normal discourse and would be unknown to most people”. Words in this band may be technical terms and part of specialised fields, such as *scintillometer*, *geogenic*, *smother-kiln* and *unwigged*. In my data from Italian we find words like *riverso* ‘a backhanded cut or stroke in fencing’ and *mezzo-relievo* ‘a relief type or technique in sculpturing’. Japanese loans in this band include *yuzu* ‘fruit in the citrus family’, *mokume gane* ‘art technique’ and *Shotokan* ‘one of the five main styles of karate’.

The words in band 3 are still not commonly found in general text but are also not considered “overly opaque or obscure”, occurring 0.01 to 0.1 times per million words. These can range from technical words like *agglutinative* and *recapitalize* to more colloquial ones like *teensy*, *badass* or *dirt-cheap*, and make up 20% of the words in the OED. Examples from my data include Italian loanwords like *vermicelli* ‘pasta type’, *vivace* ‘in music: brisk or lively performance’ and *al fresco* ‘in the open air; outdoors’. Examples from the Japanese data include *katana* ‘weapon of the samurai’ and *mochi* ‘rice-cake’.

Band 4 words, while “marked by much greater specificity and a wider range of register, regionality and subject domain” than those of the higher bands, are still likely to be recognisable by the average English speaker. These make up 11% of all entries, and examples from this band include *subpoena*, *intern*, *galvanize* and *lazily*. In my data, words from this band include words such as *al dente*, *ballerina* and *impresario* from Italian, and *origami*, *karate* and *ninja* from Japanese.

Words in band 5 make up 4% of the total entries, are more commonly found in everyday use occurring between 1 and 10 times per million words but will tend to be

considered “literate vocabulary” and includes technical terms and jargon which would be seen as “distinctively educated”. Examples from my data include *samurai* and *soy* from Japanese, and words like *tempo*, *solo* and *broccoli* from Italian. General examples from this band include nouns such as *assimilation*, *penchant*, *paraphrase* and *surveillance*, adjectives like *Neolithic*, *discontinuous*, *subsist*, *gravitate* and *presuppose*, adverbs *disproportionately*, *ad hoc* and *markedly*.

Band 6 holds about 1% of all entries in the OED, which occur between 10 and 100 times per million words. The band includes a wide range of descriptive adjectives and common nouns referring to both the material and the abstract, such as *machine*, *desert*, *horse*, *stress*, *headache*, and *career*. Other examples of words belonging in this band are the basic colours, words relating to geographical origin (*Scottish*, *Italian*, *Asian* etc.), as well as major religions, political systems or ideologies (*Christianity*, *Islam*, *democratic*, *socialist* etc.). There are no Japanese loanwords in this band, but from Italian we find words like *opera*, *scenario* and *studio*.

None of the loanwords in my data are listed as belonging to the two most frequent bands, band 7 and 8. Band 7 words make up about 0.18% of entries in the OED and occur between 100 and 1000 times per million words. This band is characterised by basic everyday words such as nouns denoting body parts, measurements of time, people and the physical world (*woman*, *person*, *tree*, *food*, *water* and *house*) as well as common abstract terms like *point*, *part*, *thing*, *quality* etc. The most frequent words are found in band 8 and occur more than 1000 times per million words. They only amount to 0.02% of all OED entries, and include prepositions, determiners, auxiliary and modal verbs, the verbs *be* and *have*, and a single noun; *time*.

3.4 Factors for Analysis

The factors for the linguistic analysis that I will carry out in this project have been borrowed in part from the linguistic analysis carried out by Pinnavaia (2001), and additionally I have added adaptations in stress pattern (following Fournier (2018) and Svensson (2004)) as a category, as well as considering frequency, based on the OED frequency bands detailed in the previous section. My data will be analysed and

discussed in chapter 6, divided into early (6.1) and later loans (6.2) and treated thematically within these sections by the following factors:

Phonological adaptations

This category will not aim for a detailed look at the correspondences between each individual source language phoneme and their counterparts in the borrowing language, but rather aims to find certain trends and comment on certain potential ‘problem areas’ as outlined in chapter 5, such as unique phonemes of the source language which are not found in the phonology of English.

Orthographic adaptations

For this category, I will look at variant spellings over time recorded in my data, to find trends in the rendition of certain phonemes, as well as discuss the differences caused by adaptation through romanisation systems in the case of Japanese. Loans can either preserve the original source language spelling or adapt to suit English grapheme conventions.

Morphosyntactic adaptations

As the loan words in my data are overwhelmingly nouns, this category is mostly limited to finding trends in plural marking. Loans may either preserve the plural marking of the source language, apply English plural -s, or form unmarked plurals. One issue here is that the majority of the tokens are not marked with inflectional information in the OED. This limits the possibilities for observing overall trends, but patterns may still emerge in the data that is available.

Stress pattern

The stress patterns of the loanwords can either be preserved (except for Japanese loans which are not stressed in the source language), adapted to a more common Germanic stress pattern, or, as Fournier (2018: 13) hypothesises with the “Italian rule”, adopt a foreign stress pattern of penultimate stress.

Frequency

The loanwords in my corpus have all been tagged with the frequency band allocated to them in the OED, which enables analyses of the relationship between frequency and the levels of nativization. Based on the descriptions of the different frequency bands, I

propose that the OED bands can be used as a rough scale for the ‘core-periphery’ organisation of the lexicon used by Friesner (2009) (section 1.3), with band 8 being at the innermost core of the lexicon of English and the lower bands increasingly peripheral. The viability vs. potential problems of such an approach will be tested and discussed in chapter 6 where my data is analysed.

4 Cultural analysis: history of contact

The following chapter provides an analysis and comparison of some of the cultural and contextual factors involved in the contact Italy and Japan have had with the English-speaking world historically. The contact with Italy will be covered in 4.1, starting with the early period, and followed by the more recent centuries. Section 4.2 will cover the contact with Japan in the same order. The following analysis is not an attempt at a complete or exhaustive account of the cultural and linguistic contact between the cultures in question but will serve to shed some light on similarities as well as differences between the cultural and linguistic contact and the opportunities for language transmission. Based on this, a preliminary conclusion for H2 will be drawn in 4.3.

I have divided the loanwords from Italian and Japanese into two groups based on their first attestation in English as listed in the Oxford English Dictionary (OED). The first group consists of early loans, with first attestations between 1500 and 1700. The second group includes more recent loans, with first attestations ranging from 1700 to the present day. The division of my data is made in order to facilitate a diachronic analysis of the data, with a comparison between the nativization of early loans versus the more recent loans.

There are several arguments for using 1500 as the earliest point for my loan word data collection. One factor being the changes the English language itself has undergone since the period prior to this. Including words from earlier than 1500 would mean keeping track of the differences between Middle English and early modern/modern English in addition to the changes and adaptations that are motivated by nativization, which would be beyond the scope of the current project. Another central argument for focusing on the modern period is the comparative aspect of my project. There are no attested loans from Japanese before 1500 as there was no documented linguistic contact before this time, and so there would be no possibility of comparison for that period.

4.1 Contact between Italy and the English-speaking world

As for the early linguistic contact with Italy, the data shows that there is also a scarcity of direct loans into English from Italian before 1500. Serjeantson (1935) claims there are “very few” loans from Italian before the sixteenth century (1935: 183), and that despite “a slowly increasing acquaintance with contemporary Italian literature during the 14th century” (Serjeantson 1935: 183), the direct contact between English speakers and speakers of the various Italian dialects seemed to be limited, as there are few Italian terms to be found, and according to Serjeantson “all came through French” (1935: 183). Pinnavaia (2001) finds only 20 loanwords from Italian in the 14th and 15th centuries in her project based on loans in the OED. The fact that most loans before 1500 appear to have come into English through French would make any analysis of their adaptation difficult, as they may have been partly nativized in French first, and as such, it is better to leave them out.

The context of the contact situation between the two languages is central to the understanding of why we find so many loan words from Italian in English historically. After Latin, French and Scandinavian, Italian is the language from which the highest number of loan words in English originate, although the period of contact has been shorter (Serjeantson 1935). The historical context will also help shed light on which semantic fields these loans are primarily from and provide some background to the analysis of their integration into English.

As outlined above, the time period I have limited my data to is the year 1500 to the present day. This is then also the period of language contact which is directly relevant to my project.

4.1.1 Early contact: 1500 – 1700

According to Serjeantson (1935), the earliest loans into English from Italian were of a commercial or military character. This suggests a cultural contact situation dominated by diplomatic, military and trade relations. Serjeantson (1935) particularly points out the direct trade route between England and Venice through the Flemish trade of the Flanders galleys as an important point of contact. This trade fleet stopped in England

regularly on its return from Italy⁴, and provided a connection between the two cultures, and by extension, languages (Serjeantson 1935).

Pinnavaia (2001) goes into more detail in her analysis of Italian loanwords and includes an overview of the various semantic fields of the Italian loanwords based on each half century. In addition to terminology related to commerce, war and the military, she lists “plants, man’s physical appearance and state, sensations and perceptions, actions, social groups, behaviour ... linguistic expressions, poetry, geometry, mathematics, the textile industry ... and sports” (Pinnavaia 2001: 153) as central semantic fields for the loans with first attestations in the OED in the second half of the 16th century.

4.1.1.1 *The Renaissance*

A central element in the cultural contact between England and the Italian region well into the 16th and 17th centuries is the Renaissance, a period of intellectual and artistic developments and innovation inspired by the wisdom and aesthetics of the classical world of the Greeks and Romans. The centre for this "rebirth of concepts and values from classical times" (Hunt 1999: 1) as well as the developments of new ideas, both in academia and in the arts, was Italy. Hunt (1999) explains that the exact dating of the period is a constant topic for debate among historians. According to Hunt (1999), those primarily concerned with the artistic developments of the period will point to the early 14th century Italian artist Giotto di Bonbone as an important early Renaissance figure for his frescoes, which “display ... the realism associated with the whole Renaissance movement” (Hunt 1999: 1). However, the spread of art styles and ideas was not instantaneous, and neither was that of words – the word *fresco* is among the Italian loanwords that made it into English in the 16th century, with a first attestation in the OED from 1598. Even though the Renaissance in Italy arguably began in the 14th century, the English Renaissance has been dated by some⁵ to late the 15th or early 16th to

⁴ Technically, there was no “Italy” as a nation state at that point in time, as the Italian unification did not occur until the 19th century. However, for the sake of convenience, I will use the term to refer to the areas which currently belong to Italy, and in which Italian dialects were spoken.

⁵ This dating is evident just from looking at titles of works on the period, e.g. Andrew Hadfield’s “The English Renaissance, 1500-1620” (2001) and Susan Wiseman’s “Writing metamorphosis in the English Renaissance: 1550-1700” from 2014.

the 17th century. That this roughly coincides with a period of large-scale borrowing from Italian is no coincidence.

Pinnavaia (2001) highlights the spread of humanistic ideas from Italy, and a renewed interest in classical languages which led to “a passion for Italy, considered the cradle of ancient and classical studies” (Pinnavaia 2001: 156) among the English during the Renaissance period. Direct communication between England and Italy was established through “ambassadors, men of clergy, diplomats, tradesmen and teachers” (Pinnavaia 2001: 156), which brought the customs and culture of Italy closer.

Historians like Hanlon (2000) characterize the concept of Renaissance humanism as “an intellectual predisposition that conveyed to modern men the treasure of knowledge, beauty and wisdom amassed by ancient Greece and Rome” (2000: 240). With the spread of humanistic ideas and the philosophy of science through Renaissance humanism, the period subsequently brought borrowings from Italian which “exalt the concept of erudition and knowledge ... related in particular to poetry, geometry and mathematics” (Pinnavaia 2001: 157), which were particularly fashionable subjects for scholars at the time. While much of the vocabulary of science was (and indeed still is) dominated by classical Latin, we see the appearance of Italian words such as *telescope* (OED: 1619) and *microscope* (OED: 1648), which illustrates the importance of Italian scientists in the scientific developments of the Renaissance and beyond. In the early 17th century Italy had 25 universities, which was the greatest density of any country in Europe at the time. The University at Padua in the Veneto region attracted many foreign students, including Englishmen (Simone 2003). The first institutional observatory was built in Padua, and they were also leading in medical research in Europe at the time.

A highly central figure of Italian science was Galileo Galilei (1564 – 1642), who studied at Padua, and whose observations of the positions of planets through an early telescope prototype he designed finally led to Copernicus’s theories of “the heliocentric system [being] fully demonstrated” (Hunt 1999: 79). Italy continued to be the seat of research into astronomy and the cosmos for a long time, as its workshops crafted the best glass lenses for telescopes (Hanlon 2000: 250), as illustrated by words related to glass production such as *calcar* (OED: 1662) ‘a furnace for glass-making’ and *ferretto* (OED: 1662) ‘copper calcined with brimstone, used to colour glass’.

Another central semantic area from which many Italian words and terms were borrowed during this early period is that of the arts; both fine arts and music feature with first attestations in the 16th to 17th century. Words for musical instruments, types of music, singing styles, art; painting techniques and sculpturing methods are recorded with first attestations in the OED from the mid-16th century onwards. Examples include words such as *cameo* (OED: 1561), *amoretto* (OED: 1596), *canto* (OED: 1590) and *alto* (1597). By the 17th century there is more, such as *opera* (OED: 1638), *capriccio* (OED: 1616) and *allegro* (1683) related to music, and art terms such as *pieta* (OED: 1660).

4.1.1.2 *Literary interest*

The English interest in Italy and the Italian language also included the literary sphere. As mentioned earlier, the translation of Italian works began in the 14th century, and this interest in Italian literature continued and gained further momentum after 1500. Italian writers such as Dante, Petrarch, Boccaccio, Machiavelli and Tasso cause “an explosion of interest in Italian language and literature” (Pinnavaia 2001: 165) in England, and many sought out sources for learning the language in order to read the works of such authors in their original form. According to Lawrence (2006), “language-learning manuals” (119) became highly popular in the late Elizabethan period, and while “Italian teachers arriv[ing] at court” (Pinnavaia 2001: 166) facilitate learning for the higher classes, the spread of didactic works and dictionaries made it possible for Englishmen to learn Italian even without a private tutor.

The word *scenario* is first attested in English in 1684, originally in the context of theatre, and different regions of Italy became popular “scenarios” or backdrops for literary works by English writers. Serjeantson (1935) argues that the wealth of Italian settings in British art and literature at the time serves as evidence of how fashionable everything Italian was. Shakespeare used Italy as the backdrop for several of his plays, notably *The Taming of the Shrew*, *The Two Gentlemen of Verona*, and famously *Romeo and Juliet*, which is set in the Italian city of Verona. Lawrence (2006) outlines the debate among scholars concerning the question of how familiar Shakespeare would have been with the Italian language, and despite many claiming Shakespeare only drew inspiration from Italian works through English or French translations, he argues that it is likely Shakespeare himself had some level of proficiency. *The Taming of the Shrew* even contains some dialogue in Italian, but Lawrence emphasizes that this is

“uncharacteristic” (2006: 122) of Shakespeare, despite the numerous uses of Italian settings in his plays.

But while the plays of Shakespeare may not have been a prime source of Italian loanwords, according to the etymological data of the OED, many new loanwords were first attested in translated literary and historical works by Italian authors such as Machiavelli, Castiglione and Guicciardini (Pinnavaia 2001). Poets such as Edmund Spenser and Philip Sidney also frequently introduced neologisms of Italian origin. Spenser’s *Amoretti* (1595) and *The Faerie Queen* (1596) were modelled on the sonnets of Italian poet Petrarch (Francesco Petrarca) and the Italian epic format of Aristo respectively (Pinnavaia 2001: 168). That the literary formats themselves were also influenced by Italian culture is evidenced by the introduction of the Italian prosodic term *stanza* in 1589, and about a century later *novella* (OED: 1677). The preoccupation with Italian “scenarios” and the popularity of Italian travel literature continued into the subsequent literary periods as well.

4.1.1.3 *Travel*

Serjeantson (1935) highlights the growing fashion of travel to Italy from the 16th century onwards as an important point of exposure of English speakers to the Italian language and culture. This helps contextualise the loaning of words relating to Italian culture and life, as travellers brought back the terminology they had been exposed to and used it to recount their travels at home, both privately and through published travel accounts. Pinnavaia (2001) points out that while the travellers might have been drawn to Italy primarily for the culture, they also “start to discover the real Italy” (Pinnavaia 2001: 158) and bring back words concerning the physical geography and natural world they encounter, e.g. the word *volcano* first attested in 1613 (OED). The word *bagnio* is found as early as 1583 (originally meant ‘a bathhouse’, later also ‘brothel or prison’), and we also find terms such as *contadino* ‘a name for an Italian peasant’ first attested in the OED in 1630.

4.1.1.4 *Religion and politics*

The subsequent decrease in loans from the 1700’s reflects a breakdown of diplomatic relations between Italy and England after Charles I’s accession to the throne in 1625, according to Pinnavaia (2001). King Charles I married princess Henrietta Maria, the

sister of the French king Louis XIII, which led England to start “a new stronger relationship with France which becomes the new model of perfection” (Pinnavaia 2001: 156). This led to a decrease of popularity for the Italian language in aristocratic circles, but with a slight delay in the decrease of direct loans into English, which Pinnavaia (2001) puts down to linguistic developments generally being slower than those of politics and society in general.

Another important point of cultural and political contact with Italy involved religion, as the leadership of the Roman Catholic Church was seated in Rome. However, influence from the Catholic Church is not reflected linguistically through loanwords from Italian to a significant degree, as the language of the Catholic Church was Latin. It is in any case clear that the overall period from 1500 to 1700 was dominated by the influence of the ideals and customs of Italian Renaissance and consequently a strong interest in the Italian language, which led to a high degree of linguistic transfer in the form of lexical borrowing during this period.

4.1.2 *Later contact: 1700 onwards*

The data Pinnavaia (2001) extracted from the second edition of the OED suggests a slight decrease in the number of Italian borrowings into English in the 18th century. However, Durkin (2014) points out that when compared with the overall amount of new words coined in English in that period, this apparent “gently descending curve” of loans still represents an increased ratio of Italian loans among new words in the statistics. Durkin argues that this reflects a “continuing strong cultural influence from Italy” (Durkin 2014: 370), however the semantic fields these borrowings belong to differ to some degree, and the number of loans still point to a different intensity of contact than what was the case in the Renaissance period.

4.1.2.1 *Travel: The Grand Tour to modern tourism*

According to Pinnavaia (2001), the period from 1700 onwards represents a difference in relation between English and Italian culture. While during the Renaissance Italian humanistic ideas had attracted English attention, “in this post-Renaissance era it is rather the search for new ideas that pushes the English to rekindle their interest in Italy” (Pinnavaia 2001: 160), and actively pursue Italian cultural influences for inspiration. One of the main ways in which this interest was rekindled was through travel to Italy.

In the early 18th century it became fashionable among Britons to travel in continental Europe for extended periods of time, and the large cities of Italy, especially Florence, Venice and Rome, were among the most desirable destinations: “In 1700, Italy, to the aspiring British traveller, seemed to represent the best of Europe's cultural achievements: the legacy of classical antiquity; the renaissance recovery of the arts; the splendours of baroque town planning and architecture” (Sweet 2012: 267). This fashion of travel became known as the “Grand Tour” and was considered by many young aristocrats as an essential part of one’s cultural education. The word *cicerone*, ‘a guide for travellers’, is first attested in the OED in this period (1719), and Pinnavaia (2001) points to many new words that relate to “leisure activities, food and drink” (2001: 160) being borrowed into English as a result of English travellers becoming acquainted with “the daily routines of Italian life” (Pinnavaia 2001: 160).

However, by the early 19th century, the relations between Italy and England had changed, as the continuing commercial and manufacturing growth of what was becoming the British Empire and the decline of many of the Italian city states saw a shift in the way the British perceived Italy, and led many inhabitants of “British towns and cities ... to [see] themselves as exemplars of culture and progress” (Sweet 2012: 267) rather than looking up to Italian cultural ideals. Chaney (1998) contrasts the British travellers of the 16th and 17th centuries, who found in Italy “a country whose contemporary culture was as impressive as its past and in many respects still outshone their own”, with 18th century travellers visiting a declining Italy that was more of “a museum set in a picturesque landscape” (Chaney 1998: xi).

Political tensions also influenced the situation, as the Napoleonic wars made continental Europe a less attractive travel destination. Sweet (2012) points out that “under Napoleonic rule ... the Italian states were effectively reduced to the status of French colonies; the period marked a nadir in their fortunes, following a long period of gradual decline” (2012: 268), which further reduced their attractiveness in the minds of British travellers.

Subsequent decades again saw an increased interest in Italy, but in the 20th century, the rise of fascism in Italy and finally World War II reasonably led to a sharp reduction in tourism as Italy and the English-speaking world - chiefly Britain and

eventually also the US - found each other at opposing sides in the conflict. The word *fascism* was borrowed from Italian during this time and is first attested in the OED in 1919.

However, moving closer to the present day, travel to Italy is again immensely popular with Italy being the 5th most visited country by international tourists according to a rapport by the World Tourism Organization (UNWTO 2018). The attractions offered by 17th and 18th century Italy, such as splendid architecture, fine art, sculpture and music clearly continue to fascinate and draw visitors in the 21st century, and Italy currently has 49 cultural heritage sites listed on the UNESCO world heritage list (UNESCO), in addition to 5 natural heritage sites, which makes it the country in the world with the most listings. According to figures from the Italian bureau of statistics (Istat), Italy was visited by over 200 million international tourists in 2017 (Istat 2018). About 6% (25.8 million) of these were from the UK and US, and an additional 5 million came from other English-speaking countries such as Canada and Australia, which confirms that travel is still an important source for direct linguistic contact with Italian for English speakers.

4.1.2.2 *Music, Art and Science*

According to Durkin (2014), vocabulary related to music and the opera is very prominent in the 18th century and into the 19th as well. Pinnavaia (2001) argues that the “large number of musical terms of Italian origin” serves as proof for the end of “the great era of English Tudor music” (Pinnavaia 2001: 160) by the start of the 1700s. New lexical referents concerning musical instruments, such as the *oboe* (OED: 1726) and the *trombone* (OED: 1724), techniques and directions for manner of playing, such as the adverb *pianissimo* (PED: 1710), as well as the names of performers, such as the operatic *prima donna* (OED: 1754) were all dominated by borrowings from Italian. More examples include *a tempo* (OED: 1740), *andante* (OED: 1742), *aria* (OED: 1742), *arietta* (OED: 1742), and *brio* (OED: 1732).

Italian art and architecture continuing to be admired in Britain, and “wealthy houses come to be decorated with precious tapestries, framed pictures, marble sculptures and stucco ceilings” (Pinnavaia 2001: 160). The numbers of borrowings related to pictorial and sculptural art further underline how “Italian culture ... continues

to influence the English one ... principally tied to artistic and scientific matters.” (Pinnavaia 2001:159).

The Italian influence on English culture is also evident in science, and Pinnavaia (2001) points to various semantic categories within the field of science as recipients of many new borrowings after 1850, such as “biochemistry, chemistry, physics, geology and mineralogy” (Pinnavaia 2001: 161). Italian scientists continued to make important new discoveries, such as the first electric battery by Alessandro Volta (1745 – 1827) who also had the unit for electric capacity (volt) named after him.

4.1.2.3 *Italian cuisine and emigration to the US*

While Italian food today arguably is among the most popular national cuisines worldwide, this was not always the case if we are to judge by linguistic evidence. Serjeantson (1935) states that “words for food were never borrowed in large numbers from Italian” (1935: 190), and culinary vocabulary indeed does not make up a large semantic group among early borrowings from Italian. However, from the early 19th century this semantic category shows “a considerable increase” (Durkin 2014: 371), and by the 20th century it is the dominating group among new borrowings, and prove that “especially as far as eating habits are concerned” (Pinnavaia 2001: 162), Italian is still exerting a considerable influence on the English lexis.

As mentioned earlier, diplomatic relations between Italy and the English-speaking world were damaged by the events of World War II, and Helstosky (2008) suggests that Italian food, and the pizza specifically, became “a kind of edible good-will ambassador, repairing fractured relations between the United States and Italy” (2008: 57) in the decades following. Italian cuisine to a large degree became popularised on a global scale through the United States. Helstosky (2008) claims that the pizza went “from being strictly Neapolitan to being Italian-American and then becoming Italian” (2008: 11), and the word *pizza* being first attested in English as early as 1825 (OED) arguably supports this. While some historians claim the pizza’s rise to popularity in the US was in part due to returning soldiers from the war in Europe “nostalgically patronizing Italian-American restaurants” (Helstosky 2008: 56), it was also down to the people who ran those restaurants and cooked the food. For Italian immigrants in the US, traditional food such as pizza became a way to maintain a connection with their

homeland as well as making a living. Although the word *pizzeria* is first attested in the OED in 1901, it was not until a few decades later that “family-friendly” Italian restaurants in the US gained more general popularity outside of areas mainly populated by immigrants (Helstotsky 2008). This also arguably had the potential of facilitating linguistic contact between English-speaking Americans and Italian speaking immigrants, although it must be pointed out that these immigrants most likely would be familiar with English as well.

Regardless of the English proficiency of immigrants in the 1950’s and 60s however, the introduction of more than 4 million Italians to the US between 1880 and 1920 (Cavaioli 2008: 214) certainly must have represented a linguistic contact situation where many English speakers were exposed to Italian vocabulary. More examples of culinary terms which were first attested in the OED during this period of large-scale immigration include *pepperoni* (1888), *tagliatelle* (1899) *mascarpone* (1903), and *focaccia* (1905). Interjections such as *scusi* ‘excuse me’ are attested (OED: 1919) in colloquial speech, which might point to the integration of Italian vocabulary in the English variety which developed among Italian immigrants, and potentially among other English speakers they were in contact with as well.

4.2 Contact between Japan and the English-speaking world

The Japanese loans collected from the OED will, similarly to the Italian loans, be analysed in two groups according to their first attestations, with a divide between early loans: 1500 – 1700 and later loans: 1700 to the present day. While the loans are divided this way for the analysis, certain historical facts of the contact situation between Japan and English-speaking countries necessitate a slightly different approach in dividing up the treatment of context. I have divided it into early and later based around the period of Isolationism, as this strongly defines the opportunities for linguistic contact between the two languages.

The contact situation between English and Japanese differs in some central ways to that between Italian and English. Historically, the contact has taken place over a much shorter span of time, and early contact was limited by factors such as the significant geographical distance and specific political measures put in place by the

Japanese government that limited the possibilities of trade, diplomacy and cultural exchange. However, there are several aspects that bear similarities, namely the strong fascination on the part of English speakers with the nature, material arts and cuisine of the source culture.

It appears to be a fact that “[m]ost of the Japanese words that have entered the English language relate specifically to Japanese culture, food, and history” (Delahunty 2008: 174). Considering this, it is both interesting and useful to consider the types of current and historical contact situations between English speaking cultures and Japanese culture in order to understand the circumstances under which such culturally specific terms were borrowed into English and how this may affect how they are integrated into the English language on a structural level.

While the borrowings from Japanese that have undergone a semantic shift or widening are rare, there are examples of words which are now used in situations where they are not referring to a concept directly related to Japan or an element of Japanese culture or nature, such as *tsunami* (OED: 1897). This loanword has become the preferred term, in both scientific and colloquial contexts, for a natural phenomenon which is in no way limited to a specifically Japanese context. Another example is *tycoon* which is now most often used in the sense of ‘an important or dominant person, esp. in business or politics; a magnate’ (OED: 1857). But even for words like these the historical context of language contact is relevant in understanding how and why they were first borrowed.

4.2.1 Early contact

The early linguistic contact between English and Japanese was limited, due in part to the significant geographical distance and considerable travel time. Very little direct contact with Japanese is recorded in the 16th century – Durkin maintains that “apart from proper names, a very few items in the sixteenth century could reflect direct contact” (2014: 396). One of the few items attested this early is *bonze* ‘Buddhist priest’ (OED: 1577), but according to the OED online that word was first brought to Europe through Portuguese, in the letters of St Francis Xavier, and may arguably not reflect

direct contact at this early stage of history. The first attestations of loanwords from Japanese, including *bonze*, and *Kuge* ‘a court noble at the Imperial Court at Kyoto’, are found in a translation of a travel account by the Italian Peter Martyr; *The history of traualye in the West and East Indies* translated into English by Richard Willes and Richard Eden and published in 1577. The letters of St. Francis Xavier and writings of Peter Martyr are representative of the earliest recorded linguistic contact between Europe and Japan, as it was mostly the travel writings of traders, missionaries and explorers like them that reached the European public at the time.

In the 17th century the number of borrowings increase, and a total of 22 words are first attested in the OED for this period. These signal a different contact situation, as there is historical evidence confirming direct contact from this time.

4.2.1.1 Trade

In 1613 the East India Company established a trading post in the port town of Hirado on the western coast of the island of Kyushu, and the direct contact from this era is exceedingly well documented, as the East India Company traders left ample records, including personal diaries as well as official documentation such as ship’s logs, and accounts and bills detailing their business dealings (Clulow 2013). Due in part to a series of unwise investments⁶ and a location too far removed from the trading centres and politically important cities of the Japanese mainland such as Edo, Kyoto or Osaka, the trading post failed to make a profit and was shut down in 1623, a mere decade after starting up. Kaislaniemi (2017) explores the linguistic environment and potential language skills of the East India Company merchants who lived at the trading post, and concludes that despite the use of interpreters, they did possess some proficiency in the “local vernacular” of Japanese themselves as well. Their letters to the homeland “are full of loanwords and borrowed phrases from foreign languages” (Kaislaniemi 2017: 59)⁷, which suggests that these traders not only relied on interpreters and the use of lingua franca languages such as Malay or Portuguese⁸, but integrated elements of the

⁶ The British traders appeared to be continuously scammed by various locals who claimed to have useful connections in the court of Japan as well as the Ming dynasty of China (Kaislaniemi 2017).

⁷ Quantitative analysis reveals the letters of the EIC traders in Japan to contain 44.8 foreign words per 10 000 words (66% of which are Japanese), compared to 8.4/10 000 foreign words in general for The Corpus of Early English Correspondence (Kaislaniemi 2017:74).

⁸ These were the most common lingua franca languages used in the “East Indies” i.e. East Asia, as Portuguese traders had been travelling to these areas for decades before the English began setting out.

foreign languages (in this case Japanese specifically) into their English vocabulary, and subsequently helped spread such terms through their communications. However, as Durkin points out “it seems very unlikely that any of [the words] were known to anything more than a very small circle of English speakers in the early 17th century, nor that they have shown a continuous history in English from this period onwards” (Durkin 2014: 396). It does indeed seem like most of the early attestations of these words are best described as isolated examples. They may illustrate the linguistic behaviour of the East India Company traders in Japan but considering the lack of any record of general usage, it is likely that they have been limited to the immediate social circle of these traders rather than having spread to the general public to any significant degree.⁹

When the East India Company abandoned their trading post in 1623 “[t]he result was the virtual suspension of English contact with Japan until the nineteenth century” (Clulow 2013: 207). Besides the failure of the East India Company in establishing themselves in the Japanese trade market, there are also specific political and diplomatic reasons for the loss of contact which affected more than just trade.

4.2.1.2 Japanese Isolationism (1639 – 1853)

The single most significant reason for the scarcity of direct contact in the 17th to early 19th century was the Japanese policy of isolationism. The policy was put in place during the reign of Iemitsu, the third Shogun of the Tokugawa Bakufu (the Japanese government) in the 1630’s (Sansom 1978), and was not lifted until two centuries later due to strong international pressure from Russia, Britain and eventually the USA. The policy meant that between 1639 and 1853 the country was almost completely closed to foreign trade and travel, with only a few strictly limited exceptions. Sansom (1978) points out that in retrospect, the Bakufu were clearly aiming at “exercising full control over all aspects of the national life, economic as well as social and moral” (Sansom 1978: 43) through the policy, in particular controlling and securing a monopoly on the limited foreign trade they allowed. However, Sansom also points out that many of the prohibitions were directly related to the anti-Christian policy of the Bakufu, and their efforts to forbid the practice of Christianity and stop missionary activity, as the

⁹ This is why first attestations in the OED can sometimes be misleading, as they may (as is the case here) not be indicative of when the word actually came into any kind of regular usage.

Christian doctrine was seen as “incompatible with the feudal principles upon which the power of the Bakufu was based” (Sansom 1978: 42).

Considering that the writings of European missionaries had been among the main sources on Japanese culture and language available in England and Europe generally, this subsequently led to a significant drop in linguistic and cultural exposure. Only 4 new borrowings are attested in the OED with first attestations in the half century following the implementation of the policy (1650 – 1700): *Dairi* ‘properly the palace or court of the Mikado: also a respectful mode of speaking of the mikado or emperor’ (OED: 1662), *moxa* ‘a soft wool prepared with down from the young leaves of any of various Asiatic plants; used as treatment in alternative medicine or as cauterizing agent’ (OED: 1675), *sake* ‘rice-based alcohol’ (OED: 1687) and *soy* ‘a sauce made from soybeans’ (OED: 1696). In addition to strict limitations on foreign missionaries, traders and travellers coming in to Japan, it was largely illegal for Japanese nationals to travel abroad. Any Japanese subject who had resided abroad for an extended period would upon returning be put to death (Sansom 1978). In practice this policy meant that any direct contact between Japanese nationals residing in Japan and English speakers or English-speaking communities was extremely limited during this time, and any linguistic transfer thus also understandably lacking. Despite this, 49 new borrowings have their first attestations in the OED during the late 17th to 18th century, such as *samurai* (1727), *Shinto* (1727) ‘the native religious system of Japan’ and *Zen* (1727) ‘A Japanese school of Mahayana Buddhism’. However, of the loanwords with first attestations in this period, the majority are limited to a specific source: the German traveller Engelbert Kaempfer’s *History of Japan*, which was translated into English by J. G. Scheuchzer and published in 1727. A total of 43 tokens have first attestations from his work. Kaempfer (1651 – 1716) was a German naturalist and physician, who travelled extensively in Asia with Dutch East Indies Company, where he worked as a surgeon (Haberland 2012). According to Haberland (2012), Kaempfer’s scientific approach and attention to detail made his work on describing foreign cultures influential in European discourse well into the 19th century, and his *History of Japan* was the main source of Western knowledge of Japan during the centuries of Isolationism.

Durkin (2014) stresses that it “should be noted” that many of the first attestations from before the mid nineteenth century are “likely to show isolated first uses, with a

long gap following before the period of fairly continuous use in English” (Durkin 2014: 397) and in light of the isolationist policies and scarcity of documented direct contact during this time, this is an important fact to keep in mind. It appears that most of the first attestations from this time are to be found in translations of other Europeans’ travel accounts. In addition to Kaempfer’s work from 1727, another example of this is the word *koto* ‘string instrument’, whose first listing in the OED in 1795 is taken from the translation of the travel account of Swedish botanist and surgeon Carl Peter Thunberg (OED s. v. *koto*). Like Kaempfer, he was travelling with the Dutch merchants, who were among the few foreigners allowed limited trade with Japan at the time, from their outpost on the landfill island of Dejima in the Nagasaki harbour (Gordon 2003: 17).

The beginning of the end of the Tokugawa isolationist policy was arguably marked by the arrival of gunboats led by the American Commodore Matthew Perry in 1853 with an “offer” that essentially boiled down to “agree[ing] to trade in peace, or suffer the consequences of war” (Gordon 2003: 49). By 1858 the Bakufu government leaders had signed a trade treaty with the United States, and subsequently the Dutch, the Russians, the British, and the French as well (Duus 1998), and Japan was now included in the growing global network of trade. The commercial contact that followed, while in no way unanimously supported by the various feudal lords (daimios) of Japan, led to a considerable increase in the social and linguistic contact between Japanese and English, especially through the increasingly close ties with the United States.

4.2.2 Later contact

New borrowings relating to political leadership with first attestations in the 19th century reflect the renewed diplomatic contact with Japan after the centuries of isolationism. Examples of such terms include *daimio* ‘title of the chief territorial nobles of Japan’ (OED: 1839), *tycoon* ‘(orig.) the title by which the shogun of Japan was described to foreigners’ (OED: 1857), and *jito* (*gito*) ‘in the Japanese feudal system: a military land steward’ (OED: 1832). But words from the semantic categories of art and material culture, nature, religion and clothing also feature with first attestations in the second half of the 1800s and are testament to an increasing interest among English speakers in a culture which had been almost completely out of reach before. Examples include *tofu*

(OED: 1880), *kimono* (OED: 1886), *geisha* (OED: 1887) and *sushi* (OED: 1893). As may be expected given the new possibility for travel, trade and other forms of direct contact, the number of loans generally introduced into English from Japanese increases exponentially: from 21 borrowings attested in the first half of the 19th century to a total of 184 from 1851 – 1900.

Durkin outlines that some important direct contact situations are those that involve English speakers living in or in close proximity to communities that are majority Japanese-speaking, which in more recent centuries include traders, the personnel of American military bases in the post-war period and more recently “ex-pat teachers” or business people in various fields, especially in tech companies, all living in Japan. He also points out Japanese emigrant or ex-pat communities in English speaking communities (Durkin 2014). These are areas that may involve some degree of bilingual contact, but Durkin (2014) argues that the main motivation for borrowing from Japanese throughout history has been Western interest in Japanese history, material culture, art and cuisine.

4.2.2.1 World wars and American Military Bases in Japan

Although the trade relationship which began as a result of Commander Perry’s excursion to Japan in 1853 had been maintained between the United States and Japan, the diplomatic relations between the two countries were not always the best. During first world war, Japan were not directly involved in military operations to any significant degree and were content with reaping the economic benefits of the European powers withdrawing from the markets of China and India (Duus 1998). However, in subsequent decades Japan engaged in multiple wars to procure new colonial territories in Taiwan, southern Sakhalin, China and Korea. As a consequence of the isolationist policies which had been in place for so long, the Japanese diaspora was not large, but the Japanese who lived abroad at the turn of the century were “mainly in the continental United States or the Hawaiian islands where they worked as field hands” (Duus 1998: 201), which arguably presented some opportunity for limited linguistic contact between Japanese and English. However, Japanese immigration to the US at the start of the 20th century became an issue as the Japanese faced increasing discrimination in the US, particularly in California, and immigration shifted towards the new Asian territories.

The increasingly expansionist tendencies of Japan and the increasing militarism of many aspects of Japanese society reflected a rise of fascism “that seemed parallel to developments in Italy and Germany” (Duus 1998: 214), and the relationship with the US worsened as the US disapproved of Japanese military aggression in China. The second world war saw tensions turn to outright war, as the attack on Pearl Harbour in 1941 marked the start of armed conflict, which eventually led the US to drop atomic bombs over the cities of Hiroshima and Nagasaki in 1945. The treatment of Japanese immigrants in the US in the decades directly following the second world war included mass internment, as strong anti-Japanese sentiment among the American public dominated. But despite the lack of positive sentiments towards Japan and Japanese culture, linguistic contact still occurred, and while the introduction of new words may not have been based on a widespread admiration of their source culture, we nonetheless find a number of words related to Japanese culture first attested in the 1940’s, one example being the realm of martial arts, with words such as *dojo* ‘a room where judo is practiced’ (OED: 1942), *basho* ‘a Sumo wrestling tournament’ (OED: 1940) and *dan* ‘a degree of proficiency in Judo’ (OED: 1941).

Some of the borrowings from this time are also directly related to the events of the war, including *honcho* ‘leader of a small group or squad’ (OED: 1945) which according to the OED was originally used among American prisoners of war in Japan during WWII, and *kamikaze* which was recorded in 1945 with reference to Japanese suicide aircraft attacks. After the end of the war, the US briefly took over Japanese military bases as well as establishing numerous new ones. Even though their existence have also been the source of some tensions¹⁰, the current “American military bases in Japan form the bedrock of the political relationship between Japan and the United States” (Karan 2005: 305). The interactions between English speaking American military personnel and Japanese staff and local communities surrounding military bases also presents an important direct linguistic contact situation. The American military presence in Japan is reflected by military slang words such as *skosh* ‘a little; a small amount’ (OED: 1959, from Japanese *sukoshi*) and *hoochie* ‘shelter or dwelling’ (OED: 1952, from Japanese *uchi* ‘dwelling’), which though they originally belonged to a

¹⁰ Reports of violent crime perpetrated by stationed US military personnel contributed to tension, especially on Okinawa (Karan 2005).

specific jargon eventually spread and made it into English vocabulary (particularly American English) on a more general basis.

4.2.2.2 *Japanese cuisine*

Words for different culinary items were part of the early loans from Japanese (*bento* in 1616, *soy* in 1687 and *sake* in 1696) but have increased greatly in later years, as globalisation has changed the way food is sold and has made the sharing of various ethnic and national cuisines possible world-wide. This has also had a linguistic impact, as the names for new dishes, ingredients and food preparation techniques are imported along with their referents. Japanese food has become increasingly popular in the west the last century, as evidenced by borrowings such as *tempura* ‘dish consisting of prawn, shrimp, or white fish, and often vegetables, coated in batter and deep-fried’ (OED 1920), *ramen* ‘noodle dish’ (OED 1967), *teriyaki* ‘mixture of soy sauce and various other flavourings’ (OED 1961), *edamame* ‘fresh green soybeans seasoned with salt, as an appetizer or snack’ (OED 1951), *gyoza* ‘a crescent-shaped dumpling of thin pastry dough’ (OED 1965) and *yakisoba* ‘dish consisting of fried wheat-flour noodles with vegetables in thick sauce’ (1957).

4.2.2.3 *Business and technology*

Japan in the second half of the 20th century recovered from the world war and grew to become a leading economic power. By the 1980’s Japan had become one of the world’s biggest economies, which led to the import of both products and terms related to Japanese business culture. Delahunty (2008) outlines how a number of Japanese business terms have become familiar in the English-speaking world since the 1980’s, and lists examples like *kanban* ‘just-in-time manufacturing system evolved in Japan’ (Delahunty 2008: 174), *kaizen*, which Delahunty describes as “a Japanese business philosophy of continuous improvement of working practices and personal efficiency” (2008: 174). A *keiretsu* is ‘a conglomeration of closely associated Japanese companies linked by cross-shareholdings’. These words are mostly terms with positive connotations, but there is also the term *Karoshi*, which paints a rather bleaker picture of Japanese business culture with the meaning ‘death caused by overwork’ (Delahunty 2008). There are also words linked to financial crime, such as *sokaiya* ‘share-holder who makes threats and blackmails a company’ (OED: 1971).

4.2.2.4 Japanese Popular Culture: “Cool Japan”

In recent decades, the cultural concept of 'Cool Japan' has emerged, as cultural exports such as anime, manga, TV-dramas and video games have spread and gained passionate fans in the West (Hashimoto 2018, Groot 2018). Groot (2018) outlines how turning the success of games, manga, fashion and other cultural products into "soft power"¹¹ has been a conscious promotion strategy on the part of the Japanese government. However, he goes on to question the effectiveness of this strategy and claims the attraction of young audiences globally to Japanese popular culture lacks depth. He claims the Japanese government has failed "to engage in the substantial issues" (Groot 2018: 18) and suggests that as the focus on cultural exports still has not "developed to become about important or universal political/moral values" they are failing to attract a lasting commitment to Japan among people "who might go on to champion Japanese values or Japan itself" (Groot 2018: 17), and are rather attracting mere fans of the cultural products.

Hashimoto (2018) further argues that the Japanese language has not been emphasised among the cultural exports. Whether that owes to a Japanese perception of their own language as too difficult or if it represents an underlying insecurity in the language is disputed, but Hashimoto (2018) maintains that the Japanese government has not been proactive in promoting the Japanese language along with their other cultural exports. Based on such claims it is tempting to speculate in whether there would be more Japanese loanwords coming into English if the Japanese had been more committed to promoting and fostering foreign learning and use of the language, beyond certain Japanese terms in connection with their cultural products.

But while scholars such as Groot question the long-term success of this policy in increasing the "soft power" and subsequent political influence of Japan world-wide, the existence of new borrowings related to these cultural imports are testament to a considerable popularity in English speaking countries the last few decades and up to the present. Examples include *anime* (1985), *manga* (1961), *hentai* (1990), *emoji* (1997), *kawaii* (1965) and *karaoke* (1977). Durkin (2014) points out that Japanese ranks highly

¹¹ "Soft power" is a concept first outlined by Joseph Nye (1992) and refers to a power of political influence based on admiration of the values and culture a country stands for rather than coercion through threats of economic sanctions or military power (Groot 2018).

among donor languages for borrowings in modern English in the OED3 data, and also highlights the fact that the borrowings are from a relatively short period of time, compared to Italian for instance.

Groot, nonetheless warns that "affection can be fickle" (2018: 16), and that enormous worldwide successes such as what Nintendo enjoyed with Pokémon Go in 2016, can decline just as fast, as they are "subject to the whims of consumers" (Groot 2018: 17). He suggests that Japan's cultural exports have "peaked" and are now struggling to regain momentum. Whether this prognosis will hold or not is unclear, but if it does that could mean a decrease in the amount of borrowings from Japanese in the decades to come.

4.3 Preliminary conclusions

As discussed in 4.1, Italy was a culturally dominant power in the Renaissance world, and the contact situation was strengthened by several factors: geographical proximity, the status and prestige of the language and culture of Italy as well as cultural exports. Interest in the culture, which was considered superior in many ways, led to interest in the language, and many attempted to learn Italian in order to get closer to the culture and elevate their own social status through association.

The cultural exports and some interest in the culture are both factors central to the introduction of Japanese loans as well, but the associations of social status and prestige are not comparable to the ones enjoyed by Italian. The early contact with Italy was dominated by an English public who idolised Italian civilisation from high art and music through science to the court culture, while early contact with Japan was limited to mild interest in an 'exotic' foreign culture. The import of material culture and art, martial arts and other non-material culture concepts used to describe the culture and society of Japan through direct contact was further halted by the policy of isolationism which spanned several centuries. Even after the country was opened to foreigners, the level of interest and linguistic contact was not on a comparable level to that observed with Italian in the earlier centuries.

Japan has not held the same dominant cultural position in relation to the English-speaking world, and although one cannot rule out the possibility that it might gain such significance in the future, even the later surge of interest in its cultural exports has not

yet reached a level comparable to Italian. In the early period, geographical distance was also a factor that limited contact between England and Japan. Globalisation has brought a geographically more distant culture such as the Japanese closer, but the attention is shared by many other cultures as well. Although Japanese now ranks high among source languages for new borrowings in modern English (Durkin 2014: 395), the contact is not nearly as dominating as the contact with Italian was during the Renaissance and continued to be in subsequent centuries.

As outlined in 2.4, theories of language contact stress the importance of intensity and duration of contact in the type of linguistic content being transmitted in a contact situation. Both the contact situation with Italian and Japanese are considerably more casual and less intense than what English was exposed to with French or Scandinavian in earlier centuries, and thus predictably primarily concerns borrowing from open categories such as nouns and adjectives. However, further stratification is arguably observable between the two, where the contact with Italy and the Italian language ranks considerably higher than contact with Japanese, particularly in the early period.

On the other hand, the limited contact with Japan also meant that the few contact points had more influence over the linguistic transmission. While the English nobility and aristocracy had the opportunity to learn Italian for themselves from language tutors and an increasing amount of literature, the exposure to Japanese was limited to the information available through a handful of sources. Among the loanwords from Japanese, 43 tokens were first attested in 1727, and these first attestations are all from Kaempfer's *History of Japan* (introduced in 4.2.1.2). Although one might argue that the borrowing process requires a level of propagation of new loans in the language and is not the direct result of the actions of a single speaker, Kaempfer nonetheless stands as an excellent example of how important a single person can in fact be in the borrowing process. As a native German speaker, Kaempfer could not easily distinguish between certain phonemes in Japanese as he was not fluent himself and they were not consistent with his native phonology. This opened up for possible transmission errors, and the most obvious one is observed in the word *gingko* (Michel 2005). The pronunciation of the original word is best represented in the Latin alphabet through the sequence *gi-n-ki-*

ya-u (Michel 2005: 4)¹². However, a simple recording mistake resulted in a misspelling which led the source word to be reproduced as /'gɪŋkəʊ/ (OED: RP), as conditioned by Kaempfer's transliteration of the word using the spelling <gingko>, instead of more phonetically accurate renderings like *ginkjo* or *ginkio*.

This example illustrates well the limitations of the early linguistic contact with Japanese, and the consequences the lack of knowledge of the source language had in a borrowing situation. It further highlights the importance of a select few sources in a situation where the direct linguistic contact was limited. There were no opportunities to rectify such transmission errors once Kaempfer had left Japan, as bilinguals were not involved to any documentable degree in Europe. In terms of loan phonology theories, this serves as further evidence of the validity of the perception model, which does not assume that the active 'borrowers' necessarily have any knowledge of the source language system. It further shows the importance of orthography and the possible consequences of orthographic effect in loanword transmission which will be discussed further in the linguistic analysis in chapter 6.

Based on the analysis of the cultural contact and the differences in the contact situation, and as illustrated through the specific example above, a preliminary conclusion is that differences in the contact situation could be a factor affecting nativization processes, as predicted by my hypothesis (H2).

The concrete differences and similarities in the linguistic adaptation of material will be analysed and discussed in chapter 6, following the outline of typological features of the two source languages in chapter 5.

¹² Based on transliteration of the pronunciation of the original Chinese characters using Japanese syllabary characters.

5 Linguistic features

The following section provides an outline of some main features of Italian and Japanese phonology, orthography and morphology, which is necessary in order to understand where these languages differ from the English systems. This further helps provide a basis for understanding how and why loanwords from these languages are adapted in the nativization into English. This chapter provides a general overview of the phonemes, syllable structure, stress pattern and orthographic conventions of Italian in 5.1, followed by Japanese in 5.2. More detail is added, such as the phonetic realisations of certain phonemes, in the cases where differences are deemed particularly relevant to the adaptation of loans into the English phonological and orthographic systems.

5.1 Italian features

5.1.1 Phonemes of Italian

The following phonemes are representative of standard Italian. Various regional dialectal differences are not included for the sake of simplicity. My inventories of vowel and consonant phonemes are adapted from Bertinetto and Loporcaro (2005) and cross-checked with Krämer (2009) and the official IPA chart, particularly for the vowels.

5.1.1.1 Vowels

While English (RP) has 12 monophthongs including long and short ones, the standard Italian system has 7 vowels (Krämer 2009). Italian does not have phonetic length distinction. The Italian vowels are listed in Table 1 below:

Table 5.1: Italian Vowels

<i>Type</i>	<i>Phonemes</i>
Vowels	/i/ - close front unrounded /e/ - close-mid front unrounded /ɛ/ - mid front unrounded /a/ - open central /ɔ/ - mid back rounded /o/ - close-mid back rounded

/u/ - close back rounded.

Table 1 above notably does not include any diphthongs, as diphthongs in Italian are often not analysed as segments in the same way that English diphthongs are. Krämer (2009) argues that diphthongs in Italian are combinations of segments, rather than (contrastive) mono-segments themselves. He argues that the different diphthongs of Italian represent “restrictions on the combinatorial options within higher units of organization than the segment” (2009: 52).

Most Italian “surface” diphthongs include close vowels like /i/ and /u/, which are considered phonologically very similar to ‘glides’¹³ i.e. approximants /j/ and /w/ respectively. In speech these close vowels are often produced as glides in such diphthongs, meaning they arguably constitute consonant + vowel sequences rather than typical diphthongs (vowel + vowel). If either of the vowel qualities in a pair of adjacent vowels cannot be turned into a glide i.e. /w/ or /j/, changing a “surface diphthong” (Krämer 2009) to an underlying glide + vowel construction, the vowels will be uttered with ‘hiatus’, i.e. with a break that separates the two adjacent vowels into two syllables rather than one (Hartmann and Stork 1972: 103). The word *hiatus* itself includes the phenomenon, as it is analysed as trisyllabic: hi.a.tus, as /ia/ is not a diphthong in English. In order to avoid an illicit surface diphthong from Italian in their L1, English speakers may either change high vowels into glides where possible or produce them with hiatus. A re-interpretation during transmission into English of which elements should be produced with hiatus versus as diphthongs or glide + vowel sequences may have consequences for how English speakers divide the words into syllables, which subsequently may impact stress patterns.

5.1.1.2 Consonants

Italian has a total of 24 consonants that are phonemically contrastive (Bertinetto and Loporcaro 2005), although the exact classifications differ somewhat between linguists. The approximants are, as can be understood from the discussion of diphthongs, particularly “controversial segments” (Krämer 2009: 47) and are sometimes analysed as

¹³ The term ‘glide’ is used to foreground the function of this type of sound as an intermediate, that helps the speech organs pass from the position of one speech sound to another (Hartmann and Stork 1972: 95)

phonologically identical to close vowels. However, in line with Bertinetto and Loporcaro (2005) I have included them in the following inventory of consonants. The consonants of Italian are listed in Table 5.2 below:

Table 5.2: Italian Consonants

Type	Phoneme
Plosives	/p/ – voiceless bilabial plosive
	/b/ – voiced bilabial plosive
	/t/ – voiceless dental plosive
	/d/ – voiced dental plosive
	/k/ – voiceless velar plosive
	/g/ – voiced velar plosive (symbol)
Fricatives	/f/ – voiceless labiodental fricative
	/v/ – voiced labiodental fricative
	/s/ – voiceless dental/alveolar fricative
	/z/ – voiced dental/alveolar fricative ¹⁴
	/ʃ/ – voiceless post-alveolar fricative
	/ʒ/ – voiced post-alveolar fricative
Affricates	/ts/ – voiceless dental affricate
	/dʒ/ – voiced dental affricate
	/tʃ/ – voiceless postalveolar affricate
	/dʒ/ – voiced postalveolar affricate
Nasals	/m/ – bilabial nasal
	/n/ – alveolar nasal
	/ɲ/ – palatal nasal
Rhotics	/r/ – alveolar trill
Lateral approximants	/l/ – alveolar lateral approximant
	/ʎ/ – palatal lateral approximant
Approximants	/j/ – palatal approximant
	/w/ – labiovelar approximant

¹⁴ Bertinetto and Loporcaro (2005) analyse /s/ and /z/ as dental, while others, such as Vincent (1988) classify them as alveolar.

5.1.1.3 *Notable variation*

Phonemes and phonetic realisations in Italian which may be relevant in a loanword adaptation process are particularly the palatal lateral approximant /ʎ/, palatal nasal /ɲ/, as well as the voiceless and voiced pair of dental affricates /ts/ and /dz/, as these do not exist in the English phonological system. These will presumably be re-interpreted by English L1 speakers as a phoneme perceived as representative for the acoustic signal they hear, or the phoneme they associate with a grapheme while reading.

Another particularity of Italian compared to English is the concept of geminate consonants: “A geminate can be defined phonetically as a sequence of identical articulations. It usually refers to the lengthening of consonants, e.g. Italian *donna*, which has a long [n:]” (Hartmann and Stork 1972: 93). English is likely to simply reduce these into single consonants, as it does not operate with long consonants. However, it may be that the concept of gemination will still be observable in some of the less frequent loanwords, which will be a good example of how certain constraints can be violated in the periphery of the lexicon.

5.1.2 *Syllable structure and stress pattern*

While English has a predominantly closed syllable structure (i.e. ending in a consonant) with consonant clusters of different sizes being allowed in both onset and coda positions (e.g. CVC, CVCC, CCVC¹⁵, etc.), the most frequent syllable form in Italian is the open syllable (e.g. CVCV, CVCVCV, CCVCV, etc.) (Cossu 1999). As Italian tends to disfavour closed syllables, they also have fewer consonants or consonant clusters in coda position compared to English.

Italian has variable lexical stress which appears to be largely unpredictable. Native speakers “make individual decisions ... that diverge from one another” when faced with stress placement tests of “nonce-words” (Krämer 2009: 156), as well as in actual lexemes, as evidenced by the existence of so-called “vacillating words” (Krämer 2009: 158) which have several official stress variants and are stressed differently from speaker to speaker. Italian stress can fall on any of the last three syllables of a word (antepenultimate, penultimate or final stress) or even the pre-antepenultimate syllable in

¹⁵ C stands for consonant and V stands for vowel.

rare cases. But while the system appears largely unpredictable, there are certain strong trends, such as the fact that heavy (i.e. ending with a consonant or diphthong) penultimate syllables attract stress (Krämer 2009), and that the majority of underived nouns show penultimate stress. This has led some linguists to conclude that default stress in Italian is penultimate (D’Imperio and Rosenthal 1999), but others, such as Krämer (2009) are not convinced that such a conclusion can be drawn. Fournier (2018) argues that the pattern of penultimate stress in Italian is such a well-known feature due to the high number of Italian loanwords, that it has had an impact on how loanwords from other languages besides Italian are stressed in English. Based on experiments with native speakers, Fournier (2018) concludes that English speakers “massively stress Italian words on the penultimate syllable whatever the original Italian stress patterns” (2018: 12). This suggests that despite Italian having seemingly unpredictable lexically motivated stress, English speakers tend to treat Italian words as having penultimate stress as a default.

5.1.3 Morphology: Inflection

As the majority of the loanwords in my data are nouns, the most relevant morphological process for this study is that of plural inflectional marking. Italian plural marking is sensitive to the gender of the noun. Additionally, Italian adjectives are also marked for gender and plurality in concordance with the nouns they modify. Italian has two grammatical genders; feminine and masculine. As most Italian words end in open syllables, plural marking mainly involves vowel change.

Three general principles, as outlined by Maiden and Robustelli (2007), apply to Italian plural formation. Firstly, feminine nouns ending in unstressed /a/ are replaced by word final /e/ in the plural. Secondly, all other singular nouns (and adjectives) ending in an unstressed vowel replace that vowel with /i/ to form plurals, and finally, nouns (and adjectives) ending in anything other than an unstressed vowel (i.e. a consonant or a stressed vowel) have identical singular and plural forms. Examples of these patterns: i)

*lasagna – lasagne*¹⁶, *foglia – foglie* ‘leaf’. ii) *stucco – stucchi*¹⁷ ‘fine plaster used for coating wall surfaces or moulding into architectural decorations’, *carne – carni* ‘meat’.

5.1.4 Orthography

Italian is written using the Latin alphabet, with no special letters not found in English. The spelling of Italian is known to be transparent to a high degree, i.e. very close to pronunciation in that there is a low ratio of graphemes to phonemes. Cossu (1999) describes Italian as having “a highly transparent orthography, characterised by an almost biunivocal grapheme-phoneme correspondence” (1999: 12), meaning there is a near one to one relationship between grapheme and phoneme in many cases. The vowel phonemes, “regardless of the context in which they occur” (Cossu 1999: 12) are spelled using five vowel graphemes: /e/ and /ɛ/ both represented as <e>, and /ɔ/ and /o/ as <o>, and /i/, /a/ and /u/ are represented as <i>, <a> and <u> respectively.

There is slightly more variation for the consonants, in particular for a few plosive consonants and for the affricates (e.g. /k/ and /g/; /tʃ/ and /dʒ/) (Cossu 1999: 12). In these cases, the same consonant letter renders different phonemes when combined with certain vowel letters to form complex graphemes. The simplicity and high degree of transparency of the Italian system generally may help ensure an easier transmission of the source phonemes in a borrowing situation, but it is also possible that English L1 speakers transfer some of the chaos of their native spelling system on the incoming loans, which may lead to some orthographic effects in adaptation and potential for sound change due to spelling-induced pronunciation.

5.2 Japanese features

The Japanese language is not part of the Indo-European language family and is thus ‘genetically’ more distant from English than Italian. Based on that one might expect more distinctive typological features which have no equivalent in English, and subsequently need for more adaptation in a loan transmission situation. However, in terms of phonology, Japanese makes fewer distinctions than English and has fewer phonemes in total. For syllable structure, English is comparatively freer in terms of

¹⁶ This word is taken from my loanword data (OED: 1845), where this pattern is replicated in my data, but the singular is rarely used anymore.

¹⁷ This word is taken from my loanword data (OED: 1598), where the native plural -i is still listed as a variant, but through the process of nativization, the -(e)s plural is now more common in English.

which elements can combine to form syllables. There are still some potential issues however, which the following section aims to cover.

5.2.1 *Phonemes of Japanese*

The information given will primarily concern the standard Tokyo dialect. There are other regional dialectal differences that I will not cover here. The exact phonetic descriptions of the place of articulation for the following phonemes vary somewhat between different linguists. My inventory has been compiled by merging the data found in Ohata (2004), Tsujimura (2013) and Kubozono (2015). I have not included all allophones of the different phonemes, but I will go into more detail on the phonetic level (as opposed to the phonemic level) for a few cases. Some allophones have been included as I deem them particularly relevant to the process of adaptation into English in a borrowing situation, due to L1 incompatibilities or possibilities for interference.

5.2.1.1 *Vowels*

The Japanese vowel inventory represents the most typical system of vowels cross-linguistically. Five is the most common number of vowels in languages across the world, and the following set is also the most common in such 5-vowel systems (Kubozono 2015). The following short vowels presented in Table 5.3 below also have long counterparts:

Table 5.3: Japanese Vowels

<i>Type</i>	<i>Phoneme</i>
Vowels	/i/ - close front
	/u/ - close back
	/e/ - close-mid front
	/o/ - close-mid back
	/a/ - open central

As mentioned above, the vowels in Table 5.3 all have long counterparts, and the existence of these long vowels is tied to a discussion of diphthongs, as several diphthongs in Japanese have developed over time into long monophthongs. According to Kubozono (2015), there is some dispute in the linguistic community as to which

vowel sequences constitute a diphthong as opposed to a heterosyllabic vowel sequence, i.e. a vowel sequence across a syllable boundary. Phonological considerations suggest that only three vowel sequences function as stable diphthongs in Japanese, and these are /ai/, /oi/ and /ui/. In a phenomenon known as *vowel coalescence*, vowel sequences ending in /u/ and many ending in /i/ underwent a historical shift to long vowels, as illustrated by the change in words such as /teuteu/ -> /tjoo.tjoo/ (/o:/), and /su.goi/ -> /su.gee/ (/e:/)¹⁸ (Kubozono 2015: 6, 242).

As discussed in the section on Italian diphthong formation earlier (5.1.1.1), sequences where a glide is followed by a vowel are often considered borderline diphthongs, and these sequences are found in Japanese as well. An example of such a structure is <iu>, which was realized as the diphthong /iu/ in Japanese historically, but is now realized as a glide with either a short or long vowel: /ju/ or /ju:/. The historical shift is illustrated in the Japanese word for ‘dragon’; originally /riu/, but now realized as /rju:/ (Kubozono 2015: /rjuu/). Kubozono (2015) regards glide + vowel combinations as onset-nucleus sequences rather than diphthongs, as they behave differently in Japanese phonology¹⁹.

5.2.1.2 Consonants

Japanese has 15 consonant phonemes with some allophonic variants, particularly among the fricatives and plosives, and I will cover some of these variants in detail in the next section. The Japanese consonants are presented in Table 5.4 below:

Table 5.4: Japanese Consonants

<i>Sub-type</i>	<i>Phoneme</i>
<i>Plosives</i>	/b/ - voiced bilabial plosive
	/p/ - voiceless bilabial plosive
	/d/ - voiced alveolar plosive
	/t/ - voiceless alveolar plosive
	/g/ - voiced velar plosive
	/k/ - voiceless velar plosive

¹⁸ Kubozono (2015) transcribes these long vowels as geminates, i.e. /oo/, /ee/ etc.

¹⁹ One reason being that glides do not contribute to syllable weight, which relates to the concept of ‘morae’ to be discussed in 5.2.2. Glide + vowel /ja/ is considered a ‘light’ syllable (1 mora), while a ‘true’ diphthong like /ai/ is considered a ‘heavy’ syllable (2 morae) (Kubozono 2015: 216).

<i>Fricatives</i>	/s/ - voiceless alveolar fricative /z/ - voiced alveolar fricative /h/ - voiceless glottal fricative Allophones of /h/ ²⁰ : [ϕ] - voiceless bilabial fricative (/h/ realised before /u/) [ç] - palatal/alveopalatal fricative (/h/ realised before /i/ or semivowel /j/)
<i>Nasals</i>	/m/ - bilabial nasal /n/ - alveolar nasal
<i>Liquids</i>	/r/ - alveolar flap
<i>Approximants</i>	/j/ - palatal approximant /w/ ([w ^β]) - velar approximant ²¹

5.2.1.3 Notable variation

The five vowel phonemes of Japanese are all present in the English system on the phonemic level, and thus in theory should not require much adaptation in transmission into English²². However, the exact phonetic representation of /u/ in Japanese differs from that of English as it has lost lip protrusion, and Kubozono (2015) represents this exact phonetic quality with the phonetic symbol [u] for a close back vowel without lip-protrusion. This symbol is also used in the Longman Pronunciation Dictionary (LPD) to represent “Japanese *u*” (Wells 2008: xxxiii). As the unrounded [u] is not present in English, this will have to be substituted by a different realisation in potential loanwords, presumably within the same phoneme, i.e. rounded /u/.

For the consonants, phonemes with allophonic variation of interest are mainly the liquid /r/ and the fricative /h/. The exact realisation of the liquid phoneme /r/ varies greatly, from a rhotic to a quality more like the English lateral approximant [l]. Shimizu

²⁰ These allophones are relevant to mention as neither exist in the English phonetic system, which may have consequences in reproduction and nativization.

²¹ Tsujimura (2013) uses the symbol [w] (labiovelar approximant) but points out that the exact quality of the lip movement is better represented by other phonetic symbols, presumably referring to [w^β], a velar approximant produced with lip compression, rather than lip rounding and protrusion.

and Dantsuji (1987) suggest that the Japanese /r/ phoneme is more similar to the voiced dental alveolar flap [ɾ] than to the approximant [ɹ] in American English. They further claim that “both lateral approximant [l] and flap [ɾ] [are used] as completely free variants” (Shimizu and Dantsuji 1987: 16) and attested in different positions, which is not consistent with a simple allophonic relationship with complementary distribution. In cases where the realisation is similar to English [l], it is possible that English L1 speakers may interpret it as belonging to their L1 phoneme /l/ rather than /r/. This might lead to the attestations of /l/ in Japanese loanwords in English despite /l/ and /r/ not being contrasted on a phonemic level in the source language.

It is worth noting that the cases of potential sound change due to re-interpretation of input mentioned above all assume a contact type and a transmission of words based primarily on the continuous input of acoustic signals. However, given the contact situation between English and Japanese discussed in chapter 4, it is likely that some loans from Japanese will show considerable orthographic effects, i.e. be influenced by their written representations in the process of integration into English. This is presumably especially true for older loans, as many English speakers are not likely to have heard the words uttered by native speakers and would thus not be familiar with the original acoustic signal. It may be more likely, given what we know of the early contact situation, that a trader or traveller picked up a word through exposure to Japanese native speakers and subsequently wrote it down, transliterated according to rules of spelling in their native language. The spread of these early loanwords from Japanese in English was then based on the written transliterated form of the word, rather than the original acoustic signal. Pronunciation would then be based on whatever limited knowledge the reader may have had of the source language phonology in addition to their L1 phonology. Transliteration of Japanese from their native spelling systems into the roman alphabet is known as *romanisation*, and there are several official systems of correspondence between Japanese syllabary characters and roman alphabet graphemes. These will henceforth be referred to as ‘romanisation systems’ and more detail on these systems will follow in section 5.2.4. The systematic differences in rendering the phonemes and allophonic variants of Japanese through the roman alphabet may have consequences for the pronunciation of Japanese loanwords by English speakers. Presumably, English speakers are likely to produce a phoneme that they

frequently associate with a given grapheme in reading, and if these are not consistent with the source language phoneme, this could lead to some variation and possible differences from the source etymon.

Another consonant phoneme that necessitates a closer look is /h/ and the noted allophones [ϕ] and [ç]. Not all sources are congruent with a classification of the former as an allophone of /h/, as Ladefoged and Maddieson (1996) point out that [ϕ] appears word-initially for /p/ in certain contexts in “native Sino-Japanese vocabulary” (Ladefoged and Maddieson 1996: 17). Neither [ϕ] nor [ç] are found in the English phonetic system, and the differences in how various romanisation systems choose to transcribe them point to potential variation in adaptation in English. A word such as 富士山 [ϕujisan] is most often Romanised as *Fujisan* ‘mount Fuji’, but an alternative romanisation renders the initial [ϕ] as <h>: *Huzisan*, which reflects a treatment of [ϕ] as an allophone of /h/ which this romanisation system does not specify. However, it is clear from the prevailing romanisation system (Hepburn) which uses <f> to render [ϕ], that it is not necessarily interpreted as an allophone by non-native speakers, who may opt for mapping it onto a phonetically “closer” phoneme /f/ rather than /h/.

5.2.2 Syllables and stress: mora and pitch accent

In Japanese linguistics, the concept of the *mora* has traditionally been used to analyse prosodic segments rather than the concept of the syllable. While a syllable is traditionally divided into the sub-units of onset, nucleus, and coda, where both onset and coda are optional, a mora has three possible realizations: (C)V, the first part of a geminate (i.e. long consonant) or “moraic” nasal /n/, which corresponds to a syllable-final or coda /n/ (Tsuji-mura 2013: 65-66). Generally, “light” syllables are considered equal to one mora, while heavy syllables consist of two morae. The difference between analysis using the concept of morae versus that of syllables can be seen in a word such as *Nissan*. In terms of syllables, this word is analysed as disyllabic: *Nis.san*, while Japanese speakers using the concept of morae would classify it as tetramoraic: *Ni.s.sa.n* (Otake 2015). This word thus illustrates all the possible mora formations: *ni* is CV, *s* is the first part of a geminate, *sa* is CV, and finally *n* is a moraic nasal. As the above mora options show, Japanese does not allow complex consonant clusters the way English, and to a more limited degree Italian, does. Japanese phonology does not allow consonant

clusters in syllable onsets, and for coda consonants, only the nasal /n/ or the first part of a geminate are allowed.

The difference in analysis of the prosodic arrangement of segments will probably not lead to any major transmission issues, as the phonotactics of Japanese lends itself well to be transferred into English syllables. One difference will be that English does not have geminate consonants and will be likely to analyse the first part of a geminate as part of the onset of the next syllable rather than its own unit, which might change the timing of pronunciation and the length and prominence put on segments. However, the timing of the pronunciation will necessarily change drastically, as English is a stress language and will automatically assign lexical stress on new vocabulary items to conform to the English prosodic system.

The analysis of Japanese words in terms of morae is closely tied to the way Japanese speakers mark prominence of certain segments. Unlike Italian and English, Japanese is not a stress language. The standard Tokyo dialect of Japanese has a distinctive pitch accent system where each of the morae of a word are produced either with high (H) or low (L) pitch, and words can move from high to low pitch (i.e. H.L.L) or low to high (i.e. L.H.H etc.)²³. Furthermore, one mora in a word may be accented, which will have consequences for the pitch. The accent marks the location in the word where the pitch falls, meaning the accented mora and the preceding morae all receive high pitch, while the following morae are realized with low pitch. The exception is the first mora of a word, which will always receive low pitch unless it is accented²⁴. Though they appear in some ways to be similar, the phenomenon of accentuation is not the same as word stress, and a large amount of native Japanese words are unaccented. Pitch accent is contrastive, as shown by a minimal pair such as [ka.n], which produced with a H.L pitch contour gives the lexeme meaning ‘can’ while a L.H pitch contour produces the lexeme meaning ‘sense’ (Tsujimura 2013: 85). Examples of the pitch patterns of accented words are [ko.*ko.ro]²⁵ ‘heart’, which has the pitch pattern L.H.L, and [*a.me] ‘rain’ with H.L pitch contour, which contrasts with unaccented [a.me]

²³ Full stops indicate mora boundaries.

²⁴ This is known as the Initial Lowering Rule (Tsujimura 2007: 68).

²⁵ * indicates accent is on this mora.

‘candy’, which has an L.H pitch contour (Tsujimura 2007: 68-69), in line with the Initial Lowering Rule.

The typological difference between pitch accent and stress will have consequences for the adaptation of Japanese words into English, and it differentiates Japanese loans from Italian ones, as there is no stress pattern to borrow along with the other phonological information of a word. One hypothesis for the adaptation process is that English will assign stress on accented syllables/mora where available, but this will not be possible to test in the present thesis, given the lack of accentuation information in most common Japanese dictionaries. It is also unlikely that this type of information will be transferred in a borrowing situation given the limited extent of the linguistic contact. The option hypothesised by the present thesis however, is that English will assign “Italian foreign stress”, as proposed by Fournier (2018) i.e. penultimate stress, to Japanese loans, particularly those with an open syllable structure.

5.2.3 Morphology: inflection

As my loanword inventory from Japanese mostly belongs to the part-of-speech class of noun, plural marking is the main morphological process that might be relevant in nativization in a borrowing situation. The adoption of English plural *-(e)s* versus the retention of source language plural marking can be a good clue to the extent of nativization.

However, Japanese does not have obligatory morphological marking of plurality on nouns. Plurality can be marked in Japanese through the marker *-tati*, but the use of this marker only applies to nouns denoting humans, such as in *gakusei-tati* ‘students’ (Tsujimura 2013: 127). Some linguists even claim that a better term for the particle *-tati* is ‘collectivizer’ rather than plural marker, as it can be used when referring to one member of a group (Tsujimura 2007: 198). The expected plural form of Japanese nouns is thus an unchanged/unmarked plural. And the presence of *-(e)s* plural would thus be indicative of nativization in progress.

5.2.4 Orthography

Japanese is written using three different orthographic systems concurrently: *kanji*, *katakana* and *hiragana*. The latter two are syllabary systems which represent the

possible mora combinations of Japanese phonology. They are both historically derived from simplified kanji. The term *kanji* refers to the characters based on the Chinese logographic system, which are used to write most content words such as nouns and word stems for adjectives and verbs. The Japanese government's "List of characters for daily use" issued in the 1980s contains 1,945 kanji characters (Shibatani 2017). They are combined with *hiragana* which is used for inflected endings for verbs and adjectives and morphological particles. Hiragana is also used for Japanese words which are not represented by a kanji character (Shibatani 2017). *Katakana* consists of more angular looking characters and is now used for emphasis (much like italics are used in the Latin alphabet), in the spelling of foreign loanwords and in business names, as well as in scientific description, e.g. botanical plant names (Shibatani 2017). In addition to these three systems, the Latin alphabet is used occasionally for initialisms and brand names as well as for romanisation of the language for foreign learners and academics.

The two main systems for romanisation of Japanese are called *Kunreishiki* and *Hepburn* (University of Tokyo 2009). Even though *Kunreishiki* is the official system taught in Japanese primary schools, the Japanese government also uses the *Hepburn* system for passports and international communications, and it is recommended for academic use. *Hepburn* is also the system most often used in English-language publications both in Japan and abroad.

The *Kunreishiki* system is easier for native Japanese speakers to learn as it provides a closer one-to-one correspondence between kana and English letters, and it is thus also better suited for linguistic descriptions of Japanese on a phonemic level. However, due to phonetic (allophonic) variation, one letter may not be the best way to convey the actual pronunciation of the sounds in different kana. This is where the *Hepburn* system may provide native English speakers, who are not familiar with the allophonic variations of Japanese, with a guide to actual pronunciation on a level closer to the phonetic. A concrete example are the kana た, ち, つ, て, which would all be written with a <t> using the *Kunreishiki* system: *ta, ti, tu, te*, while the *Hepburn* system would render these as: *ta, chi, tsu, te*, which is arguably closer to their actual phonetic quality (University of Tokyo 2009). As mentioned earlier (in 5.2.1.3), the difference in romanisation may have consequences for how loanwords are adapted in English, as these writing systems may lead to orthographic effects on the pronunciation.

6 Linguistic analysis

The following chapter provides an analysis and discussion of the loanwords recorded in the early period (1500 - 1700), followed by the later period (1700 to the present) and concludes with a section comparing the trends found in the two time periods and a discussion of my research questions and the validity of my hypotheses. The data will be analysed and discussed according to the factors listed in 3.4: frequency, morphosyntactic adaptations, phonological adaptations, orthographic adaptations, syllable structure and stress pattern. Some attention will also be given to special cases, which demonstrate adaptations (or lack thereof) that diverge from the overall patterns of the data as a whole.

6.1 Early loans: 1500 – 1700

The total number of early loanwords from Italian and Japanese combined is 332. Among these, the Italian loanwords are decidedly more numerous, with 307 tokens, while only 25 tokens originate in Japanese. In terms of broad semantic categorisation (outlined in 3.4), several of the Japanese words are related to Japanese *society*, including terms describing currency and noble titles (7 tokens). Many of the tokens describe Japanese *material culture*, including art, clothing and furnishings (5 tokens). *Culinary terms* are also well represented, with 5 tokens. Among the Italian loanwords, the semantic categories that stand out are *music* (33 tokens), *art* (26 tokens), *architecture* (15 tokens) and *culinary terms* (13 tokens).

The difference in number of tokens clearly relates to the difference in the nature of the contact situation discussed in 4.3. The contact with Italian was widespread, and Italian cultural products and customs were held in high regard and the association with Italian language and culture was an essential part of presenting a cultured image for the elite, as well as being something the public enjoyed and was curious about. By comparison, Japan was still a distant country without any cultural influence or significance on the general public in Britain. The semantic categories of the loanwords underline this fact, and the presence of several words related to currency point to the limited direct contact under the short-lived existence of the East India Company trading

post in Japan. A look at the early sources for these words in the OED confirms letters and reports from these to be the original first attestations in the majority of cases.

Another perspective that relates to the levels of contact is the part of speech categories that the loanwords belong to. For the Japanese words, all 25 tokens were nouns, which is regarded as the category most likely to be involved in borrowing in a casual contact situation. Among the loanwords from Italian however, tokens belonging to other part of speech categories were attested. As many as 41 tokens were adjectives or used adjectivally, and 12 tokens were listed as an adverb or as used adverbially. The data also included 5 verbs and 8 interjections. The presence of categories such as adverbs and interjections in particular, point to a higher degree of intensity of contact, as these are not among the categories most often involved in borrowing cross-linguistically, according to the language contact hierarchy of Matras (2007) as discussed in 2.4.

6.1.1 Frequency

The distribution of Italian and Japanese loans with first attestations between 1500 - 1700 across the OED frequency bands is presented in Table 5 below. The characteristics of the different bands were outlined in 3.3. Band 7 and 8 are not included, as there are no tokens from either language recorded in those.

Table 6.1: frequency distribution of early loans

Frequency	Italian	Japanese
<i>Band 1</i>	13	0
<i>Band 2</i>	92	6
<i>Band 3</i>	85	13
<i>Band 4</i>	70	5
<i>Band 5</i>	39	1
<i>Band 6</i>	5	0
Total tokens²⁶	307	25

As is clear from the number of tokens, there were significantly more loans imported from Italian than from Japanese during the early period, which is not surprising given

²⁶ The total includes 3 tokens from Italian where no frequency band was listed.

the contact situations outlined in chapter 4. It is interesting to note that the largest category for Japanese loans was band 3, while the largest category for Italian was band 2. While the difference in raw numbers makes it hard to draw comparisons between the two, the data from Italian seems to suggest that a high amount of Italian loans from this period have either stayed in the periphery in terms of frequency, or been relegated there as they have fallen out of use and popularity in recent decades. Table 6 below presents the frequency data as percentages to allow for easier comparison. The six frequency bands have been paired together in three categories as *low* (1&2), *middle* (3&4) and *high* (5&6). As mentioned above, there are no tokens from band 7 and 8, and thus 5 and 6 constitute the highest frequency in my data.

Table 6.2: frequency distribution of early loans: percentages

Frequency	Italian	Japanese
<i>Low</i>	34.2%	24%
<i>Middle</i>	50.4%	72%
<i>High</i>	14.3%	4%

As shown in Table 6, the percentage of the total which are found in the higher bands also underline how a decent number of early Italian loanwords are still commonly in use among speakers. 14.3% percent of the total early loans from Italian are in band 5 and 6, compared to the single token in band 5 for Japanese, which makes up a mere 4% of early loans from Japanese. The combined percentages of the middle bands 3 and 4 for the Italian loans amount to 50.4%, while for Japanese the number is 72%. This indicates that the majority of these early Japanese loans are in use with moderate frequency today. They are not considered overly obscure, as discussed in 3.3, but rather likely to be recognisable by the average English speaker. Examples include *katana*, *bento* and *miso*. For the two lowest bands the percentage is 33.9% for Italian and 24% for Japanese, which indicates that a slightly higher percentage of the words from Italian are considered obscure and rare compared to those from Japanese.

The number of tokens observed in the table above supports an analysis of a less close contact situation with Japan compared to Italy between 1500 and 1700 (as discussed in 4.3). However, it must be stressed that any conclusion about the historical

contact situation cannot base itself on frequency, as this is an entirely synchronic notion, as discussed in the methodology chapter (3.3).

6.1.2 *Phonological adaptations*

Pinnavaia (2001) divides the phonological adaptations she observed in the nativization of Italian loanwords into three processes, using terms borrowed from Gusmani (1987), and these are *approximation*, *mechanic adaptation* and *analogy*. Approximation is used to describe the process whereby speakers attempt to imitate the pronunciation of a phoneme from the source language, which is not found in their L1, by replacing it with a phonetically similar sound in their L1 system. The second process, mechanic adaptation, involves the substitution of foreign phonemes based on the rules of the L1, or the position of these phonemes “along the phonetic chain which in any language responds to its own phonotactic rules” (Pinnavaia 2001: 91). One example is the phoneme /r/, which is always voiced in Italian, but is only realised in English RP pre- or intervocalically. Through the process of mechanic adaptation, post-vocalic /r/ would thus be elided in loanwords when reproduced in RP. The last process is analogy, which refers to cases of adaptation as a result of analogy with native forms, through the “interpretation of foreign phonetic patterns ... conditioned by their spelling” (Pinnavaia 2001: 93), which can also be defined as cases of orthographic effect. An example of this found in my data, is the reinterpretation of Italian words ending in <ese> /ese/ as the English suffix <ese>, which is normally pronounced /i:z/.

The most prominent nativization trends concern the adaptation of vowels, and particularly interesting are the trends which are to be found in the adaptation of unstressed vowels in open final syllables. The two clearest trends are arguably consistent enough to be considered *rules*, rather than mere trends. One such consistent phonological adaptation (or rule) is the phenomenon of vowel breaking, or *diphthongisation*, wherein source language monophthongs become diphthongs when reproduced in English. The adaptation is observable in Italian as well as Japanese loanwords and involves the change from word final unstressed monophthong /o/ to diphthong /əʊ/ in RP variants, and /ou/ in GA variants. This process is observed in all 4

eligible tokens from early Japanese loans²⁷, and the clear majority of Italian tokens. Only 9 Italian loanwords ending in /o/ show resistance to this process by preserving /o/, and one such example is *giro* /'dʒiro/. Pinnavaia (2001) also observes this process and suggests that it might be the result of English speakers trying to maintain the original source language vowel sound, which would otherwise likely be reduced to schwa /ə/ (2001: 89), as is common in unstressed final syllables in English.

One consequence of diphthongisation is that the syllables in question gain syllable weight: light final syllables become heavy ones. It thus becomes apparent that English (as expected) does not consider the phonological implications of added syllable weight considering the concept of the mora in Japanese linguistics (discussed in 5.2.2). Light syllables count as a single mora while heavy syllables count as two, and the number of morae has consequences for the prosody of an utterance. The adaptation of loanwords from Japanese thus appears to not be sensitive to the Japanese concept of morae, and the moraic system is translated into that of syllables, in which weight does not have an effect on the number of syllables or on stress patterns.

The other trend regular enough to be considered a rule is that of final unstressed vowel reduction to schwa /ə/. This trend mainly involves word final /a/. Among the Italian tokens ending in <a>, 56 tokens reduce to /a/ to /ə/, and only 2 tokens resist this process. These are *giunta* /'dʒʊnta/ and *tartana* /tar'tana/. Their resistance can partly be explained by the limits of their usage. *Giunta* is recorded as being in *historical usage* only, while *tartana* has a synonym *tartan* which belongs to a higher frequency band, which arguably relegates both of them to the periphery of the lexicon. Only 2 early loanword tokens from Japanese end in /a/, and both (*katana*: /kə'tɑ:nə/, *moxa*: RP: /'mɒksə/, GA: /'maksə/) undergo reduction to schwa, reaffirming the trend from Italian.

For words ending in unstressed /e/, the trend is not as clear. Pinnavaia (2001) groups adaptations of word-final /e/ together with the vowel breaking of /o/, but I would argue that as there is much more variation in the adaptation processes that involve /e/, the diphthongisation process is much clearer as a main process and rule of the nativization of final /o/. Among the 4 tokens from Japanese that end in word final <e>, three different strategies of nativization are observable: diphthongisation (*saké*: GA /'sakeɪ/),

²⁷ i.e. words ending in unstressed /o/.

realisation as /i/ (*saké*: RP /'sa:ki/), /ɛ/ retained (*kuge*: /'ku:ɡɛ/). A fourth strategy is that of apocope, the “omission of one or more sounds or letters from the end of a word” (Hartmann and Stork 1972: 16), where the final <e> is lost in pronunciation. This appears to be exemplified by *bonze*: RP /bɒnz/, however, a look at the etymology of this word shows that the original etymon is *bonzu*, which means that the silent final <e> is an addition in the nativization in English, after final <u> was lost through apocope. These four processes are also observed in the Italian loanwords ending with <e> in an unstressed position. Diphthongisation is found in tokens such as *canzone*: /kant'səʊneɪ/, *marchese*: RP /mɑ:'keɪzeɪ/, and *rione*: RP /ri'əʊneɪ/. Some words display variation, such as *padrone*, which is listed with vowel breaking /pə'drəʊn(eɪ)/, but also /i/, both variants in RP: /pə'drəʊni/. Additional examples of <e> being realised as /i/ include *podere*: RP /pɒ'dɛ:ri/, *consigliere*: RP /kɒnsɪli'ɛ:ri/ and *cameriere*: RP /kɑmɛri'ɛ:ri/. Examples of words which retain a monophthong include *vivace*: /vi'vɑtʃɛ/, *sestiere*: /sesti'ɛrɛ/ and *campanile*: /kɑmpɑ'nɪle/. Further examples of loss of final <e> through apocope will be discussed in 6.1.3.

Some main trends in the adaptation of consonants on a phonological level, is the process whereby geminate consonants become single consonants. English does not operate with gemination in consonants except for in a few limited contexts, such as derived or compounded words (*bookkeeper*: RP /'bʊk,ki:pə/, is one such example). The adaptation of geminates into single consonants is regular, but a handful of tokens resist this process, and are listed with transcriptions in the OED that retain gemination. Examples include *cavetto*: /ka'vetto/, *beccafico*: /bɛkka'fɪko/ and *capriccio*: /ka'prɪtʃo/. There further appears to be a correlation between words which retain gemination and words resisting other nativization processes, such as vowel breaking. All three of the previous examples have also retained word final /o/, and other words display additional resistance to nativization, such as *sbirro* ‘an Italian police officer’. The transcription /'zɪrro/ violates several well-formedness constraints of English, in addition to gemination, the initial consonant cluster is illicit, as /z/ does not form clusters in syllable onsets in English (Harrington and Cox 2019). The expected adaptation in the onset cluster, would be a loss of voicing, producing /sp/ instead, making the full word one would expect if nativization had taken place here something like /'spɪrəʊ/. The word belonging to frequency band 2 can be seen as an explanation of why such illicit

structures have been reproduced and not undergone any adaptation. The fact that the words containing gemination all belong to the lower frequency bands (mostly band 2 and 3) appears to be in line with core-periphery theory, which states that peripheral items are “allowed” to violate constraints that apply to core vocabulary. Some of the same tokens that illustrate gemination also reproduce phonemes of Italian that are not found in the phonemic inventory of English. One example is *lazzaro*: /'laddzaro/, which appears to include the voiced dental affricate /dz/²⁸. Again, this token belongs to frequency band 2, which further supports the connection between frequency, constraint violation and the core-periphery theory.

When it comes to the adaptation of source language phonemes not found in English, one notable case in Japanese loanwords is the adaptation of the Japanese voiceless bilabial fricative phoneme [ɸ], which is found in the etymon for a loan such as *furo* ‘bathhouse’ (Hiragana: 湯 槽). The pronunciation of this word as /'fʊrəʊ/ or /'fʊ:rəʊ/ (RP variants) arguably exemplifies a sensitivity to Japanese allophonic relations, as [ɸ] is the realisation of the phoneme /h/ before /u/. However, the fact that the most common romanisation system (Hepburn) renders the hiragana symbol 湯 [ɸu] /hu/ as <fu> undoubtedly is a strong factor of influence here, and thus the use of /f/ does not necessarily indicate anything about English speakers’ sensitivity to Japanese phonetic variation.

6.1.3 Orthographic adaptations and orthographic effects

It is not easy to draw a sharp divide between orthographic adaptations and phonological ones, as the two frequently interact and influence each other. Some of the trends described here have direct correlations to pronunciation, either as a cause for a change or an effect of one, and it can be hard to determine which scenario is more likely.

One observable process involving orthographic adaptation, which was discussed in 6.1.2 is that of loss of final elements through apocope, that results in shortening. The final syllables in Japanese and Italian are usually open syllables. Some Italian loans that have undergone this type of shortening include *violin*, from the Italian etymon *violino*

²⁸ Whether this sequence is pronounced as separate sounds rather than as a single phoneme is difficult to ascertain.

and *arsenal*, from the Italian etymon *arsenale*. The loss of final <e> in *arsenale* is not surprising, given the fact that word-final <e> tends to be silent in English. The adaptation and pronunciation of final <e> is as discussed in 6.1.2 variable, and the fact that it is often silent in English orthography can lead to orthographic effects, whereby it is not realised in pronunciation, and subsequently also falls away in the orthographic representation. This would be an example of the process Coetsem (1988) calls pronunciation-induced spelling, as discussed in 2.3. This trend might also reflect the fact that the English phonotactic system prefers closed syllables to a higher degree than the source languages. Rather than a loss of graphemes, another trend involving <e> concerns alteration. For several of the tokens listed in English with word final <e>, the original etymon may have had a different vowel, such as <o> in Italian *granite* /'granɪt/ (etymon: Italian *granito*) or <a> in *nunciature*: RP /'nʌnsɪətʃʊə/, /'nʌnsɪətʃə/, GA /'nənsiə,tʃʊ(ə)r/ (Italian etymon: *nunciatura*) and *sardelle*: RP /sɑ:'dɛl/, GA /sɑr'dɛl(ə)/ (Italian etymon: *sardella*). Here the word-final /o/ and /a/ from the source etyma have been lost, and the graphemes <o> and <a> replaced with the phonologically 'empty' grapheme <e>. Again, the silent final <e> may either be due to orthographic effect or be an orthographic change due to phonological adaptation in the form of apocope. The word *nunciature*, along with *miniature* and *explicature*, also represents phonological apocope, possibly due to analogy with the pronunciation of the suffix *-ture* in other English words, after orthographical adaptation changed the final <a> in the etyma to <e>. The original Italian /tura/ is subsequently realised as RP /tʃə/ or /tʃʊə/ due to mechanic adaptation rules, and as /tʃʊ(ə)r/ or /tʃər/ in GA.

Some examples of less numerous orthographic adaptations include metathesis, as observed in *palander*: UK: /'pæləndə/, US: /'pæləndər/, from the etymon Italian *palandra*. Here the consonant cluster <dr> has been broken up by the switching of <a> and <r>. In British English this subsequently leads to loss of /r/ through mechanic adaptation, as postvocalic /r/ is not produced in speech. The word final <a> has been reduced to schwa in both the British and American variants, following the trend of unstressed /a/ reducing to schwa discussed above.

A good example of possible orthographic effect is found in *vermicelli*, which is listed with two pronunciation variants: /və:mɪ'sɛli/ and /və:mɪ'tʃɛli/. The first variant here reflects a reading of <c> as /s/, which complies with the English rule of <c>

realised as /s/ before <e, i, y> while the second option reflects the source language rule of realising <c> as /tʃ/ before <i> or <e> (Maiden and Robustelli 2007: 4). The occurrence of two variants arguably points to this word as a case of nativization in progress, where one variant resists adaptation based on the rules of the English system, while the other adapts to comply. There are also early spelling variants that suggest orthography could have been an influencing factor. The OED lists *vermechulli*, *vermicelly* (-cella), *virnizzelli*, and *vermeselly* as variants used in the 17th century, and it is possible that the use of /s/ in pronunciation was influenced by the latter two of these. Alternatively, they may simply be results of attempts to reflect the /s/ pronunciation variant orthographically. Again, it is difficult to determine which adaptation came first of the phonetic or the orthographic in cases like these, and thus difficult to conclude if the observed changes are down to orthographic effect or phonological adaptation.

Many of the Japanese words appear to have undergone a standardisation of spelling based on the romanisation systems discussed in 5.2.4, as opposed to natural adaptation in the language over time based on phonotactics and common grapheme-to-phoneme correspondences in English. Examples of some interesting variants show that the early orthographic variants do not appear to have significantly affected the pronunciation. One factor might be that these words were (as discussed in 4.2) in use only in limited circles of speakers and did not spread into wider use until much later. One might argue that the amount of time passed qualifies some of these as reborrowings. Some examples are *furo*: RP /'fʊrəʊ/, /'fʊ:rəʊ/, GA /'furoʊ/, where earlier spelling variants from the 17th century like *froo* and *fouro* do not reflect the current pronunciation, and *samisen*: RP /'sɑmɪsən/, GA /'sæmə,sən/, where several variant spellings include <sh>, such as *shamshin* attested in the 1600s. Applying English grapheme-to-phoneme conventions, the use of <sh> might lead to the appearance of /ʃ/ in pronunciation, which is not the case in the current transcriptions.

However, there are also examples of tokens which do not follow the standard set by romanisation systems, such as *itzebu*, whose Japanese etymon いちぶ according to a standardised Hepburn romanisation would be rendered as *ichibu* which is in fact attested among the historical variants. However, the current pronunciation shows clear orthographic effect from the *itzebu* spelling: /'tsɪ'bu:/. The token *moxa* is another

unusual example, as its Japanese etymon もぐさ, would be Romanised as *mogusa*. The OED theorises that devoiced [w] might have led English speakers to perceive the word as /'mɒksa/, leading to the spelling *moxa*²⁹.

The last example could be seen as a counterargument for the theory that orthographic effects have been particularly evident in Japanese loanwords. However, one could also argue the opposite, that this example in fact helps to illustrate the nuances of the borrowing situation from Japanese. While the original borrowing scenario necessarily had to involve acoustic signals, as the original adapters were unlikely to be fluent enough in written Japanese (as discussed in 4.2.1) to have picked up words from their visual representation, the later spread of these loans among the general public was heavily influenced by (perhaps even determined by) the written representation, i.e. visual sign used by the initial borrowers rather than the original acoustic signal from Japanese.

6.1.4 Morphological adaptations

As previously mentioned in 3.4, the information available on morphological plural inflection is limited to a minority of the total tokens in my data set. This fact limits the usefulness of any conclusions that might be drawn from the data on this aspect, but I will nonetheless attempt to extract some trends and discuss them. The native Italian plural marking was discussed in 5.1.3, while Japanese plurals were covered in 5.2.3.

Among the 307 early Italian loanwords in my data, 207 tokens do not have any inflectional information listed in the OED, but in addition to nouns, this number also includes several adjectives, adverbs and interjections. Of the 100 nouns with plural information specified in the OED, we find source language plural marking, variable plural marking, unchanged plurals and English plural marking. Based on the divisions that this information makes possible, it is tempting to draw up a cline of morphological nativization moving from least to most nativized through three stages: Source language plural → Variation between source and target language plural → Target language

²⁹ And later final vowel reduction to schwa.

plural. But this path is not mandatory, as it is also possible for loanwords to skip these stages and take the native *-s* plural directly.

Among the Italian loanwords with plural marking included in the listings, 29 tokens are attested with the Italian plurals *-e* or *-i* only. None of these words belong in the higher frequency bands of 5 and 6, and only four tokens belong to band 4, which arguably supports the theory of the inflectional cline proposed earlier, as it shows that words in the periphery of the lexicon in terms of frequency indeed do remain markedly foreign in regard to inflection.

The largest group, with 33 tokens, consists of entries listed with both English *-(e)s* and Italian *-e* or *-i* plural marking. This group can thus be considered one in which nativization is still visibly in progress and finds itself between the ‘source language inflection only’ and the ‘borrowing/target language inflection only’ groups on an imagined cline of morphological nativization; from none to complete. Specification of the most common versus the rarer variant is sometimes included, such as in *opera*, which is listed with plural *-s*, but also *-e* (*opere*) as a rare variant. This arguably places *opera* further along on the cline of morphological nativization than a word such as *relievo* ‘moulding or carving in which the design stands out from the surface’, where *-s* is listed as rare, while *-i* is the more common plural marker. For these specific examples, the frequency data also supports this interpretation, as *opera* belongs to band 5 while *relievo* is band 3, additionally indicating that *relievo* is a more peripheral lexical item. However, as a general trend, this unfortunately does not add up as nicely. This group includes words from every frequency band between 1 and 6, and some tokens marked with Italian plurals as more common than the English *-s* also belong to higher frequency bands than those where the opposite is the case, or where no such ranking is specified. One such example is *lira* ‘Italian currency before the Euro’ which belongs to frequency band 4, but is listed with plural *-e* being more common than *-s*.

7 tokens are listed with unchanged plural forms or variants between unchanged and *-s* or *-i* plural. Among these are four words that can function adjectivally as well as nominally: *replica*, *Piedmontese*, *Albanese* and *Lucchese*. The latter three are terms used to describe people from a specific location, or the speech of these people. They are special cases as the ending *-ese* (Italian /-ese/) has been re-interpreted as the affix /-iz/,

which is productive in English for this use. These are thus clearly further along the cline of nativization than many other words recorded with variant inflection.

My data includes 28 tokens listed with English plural *-s* or *-es* only. These could be considered at the integrated end of the nativization cline, having fully adapted morphologically to L1 plural marking. The frequency distribution among these 28 stretches from band 1 to band 6, but the majority is in the higher bands, with 13 tokens combined in band 5 and 6. Among the middle bands 3 and 4, we find 9 tokens, while there are 6 tokens belonging to the lower bands 1 and 2 combined. This suggests a (tentative) correlation between morphological nativization and the higher frequency bands.

Additionally, we find some special cases of words being loaned into English in their plural form. Pinnavaia (2001) also observed this and suggested that for some of these entries it may be a case of word final <i> being preferred on the basis that word final <e> tends to be silent in English. One example being *macaroni*, whose Italian etymon would be *macarone* in the singular, with *-i* plural inflection. However, this noun is arguably used as a non-count noun in English and does not take any further plural marking, and so the motivation for using the plural form over the singular may just as well be semantically motivated as orthographically.

As discussed in 5.2.3, Japanese does not have obligatory morphological plural marking. This means the tokens that are listed specifically as having ‘unchanged’ plural forms could be considered examples of words that have not undergone morphological nativization, while the presence of English plural *-s* or *-es* could be indicative of nativization in progress. However, it must be noted that native English nouns can also exhibit unchanged plural forms, a famous example being *sheep*, and thus the lack of *-s* inflection need not be an indication of lack of nativization either.

Among the 25 early Japanese loanwords, as many as 19 tokens do not have any plural information listed in the OED. Among the remaining 6 tokens, one is listed with English plural *-s* only, while 5 are listed with both unchanged and *-s* plurals as variants. The words are divided between frequency band 2, 3 and 4, and there do not appear to be any specific characteristics of the group that take *-s* plural as a variant in terms of frequency or presence versus absence of historical spelling variation or other factors as

compared to the words which have no information listed. Interestingly however, among these, some are listed as nativized with English *-s* plural in earlier variants, such as the words *mochi* ‘rice cake’ and *kami* ‘nature spirit’ which both have variants with *-s* recorded in the 1600s: *muchos*, *camies* and *caymies*. These historical plurals arguably present a major complication for the viability of a representation of morphological nativization as a unidirectional cline, moving from source language plural marking to L1 plural marking, as it appears these words have gone back to varying the plural marking between unmarked and *-s*. Some Italian tokens also have other plurals attested historically, such as *soldo* ‘Italian coin: formerly the twentieth part of a lira’, which was listed with the variant *souldyes* in the 1600s, but now takes Italian *-i* plural marking only. This challenges a direct assumption that use of source plurals means a word has not yet been nativized and suggests that the interpretation of the concept of nativization as a path or a cline needs some nuance. There is no way of knowing, based solely on the OED’s data, the spread of usage these earlier variants enjoyed, but the fact that they exist, still arguably disputes the idea of the path of nativization, unless we accept this path as bidirectional, meaning the levels of morphological nativization are subject to change diachronically.

6.1.5 Syllabic structure and stress pattern

Among the total 307 early Italian loanwords, 8 are monosyllabic, 73 are disyllabic, 125 are trisyllabic, 75 are tetrasyllabic and 21 are polysyllabic, i.e. with 5 or more syllables. Additionally, 5 words have a variable number of syllables recorded in the pronunciation variants listed in the OED. The monosyllabic words are not relevant when it comes to stress placement, but the distribution of stress placement on early Italian loanwords according to number of syllables is presented in Table 7 below:

Table 6.3: stress placement on early Italian loans

	Disyllabic	Trisyllabic	Tetrasyllabic	Polysyllabic	Variants
<i>Final</i>	7	13	3	-	-
<i>Penult</i>	63	66	49	10	1
<i>Antepenult</i>		41	17	5	-
<i>Preantepenult</i>			2	1	-
<i>Variant</i>	3	5	5	1	4

For the disyllabic words, stress can either fall on the penultimate (i.e. initial) syllable, the final syllable, or a variation between both across pronunciation variants. The remaining cells in the column (i.e. antepenult and preantepenult) are thus necessarily left empty. As seen in Table 7 above, the most common stress pattern was clear, as 63 out of the total 73 disyllabic words were stressed on the penultimate syllable, while stress fell on the final syllable for only 7 tokens. 3 words had variants of both stress patterns attested.

For trisyllabic words, stress can fall on either the antepenultimate syllable (the initial), the penultimate syllable, the final or a variation of the former across pronunciation variants. Table 7 above, shows that the distribution between the stress patterns was slightly more varied for this category. The antepenultimate syllable attracted stress in 41 tokens, but with 66 tokens, the majority of the 125 trisyllabic words were still stressed on the penultimate syllable. Final stress was recorded for 13 tokens, while 5 tokens showed variation. Among words with four syllables, stress can fall on the pre-antepenultimate (initial) syllable, the antepenult, the penult, the final, or a variation of these. For the 75 tokens recorded in my data, only 2 words had pre-antepenultimate stress, while 17 had antepenultimate stress. Penultimate stress is still the dominating pattern with 49 tokens. Only 3 words had stress on the final syllable, while 5 tokens had variation recorded across pronunciation variants.

The Italian polysyllabic words in my data were stressed on either the pre-antepenultimate syllable (1 token), the antepenultimate syllable (5 tokens) or the penultimate syllable (10 tokens). There were no polysyllabic words with final stress, but one token had variation in stress patterns across pronunciation variants. Among the small group of tokens with variable syllable structure, 4 also had variable stress patterns, while 1 token, *piano*, had penultimate stress both as a disyllabic (/ˈpjɑ:nəʊ/), and as a trisyllabic word (/pɪˈɑ:nəʊ/), which may be a small indicator of the preference of the penultimate stress pattern for early Italian loanwords.

The most common stress pattern across all syllabic categories was clearly penultimate stress, while the rarest pattern was pre-antepenultimate stress. Common for the three tokens with pre-antepenult stress is a classification by the OED as “rare”. They

belong to frequency band 1 and 2, a fact which further points to their place in a peripheral part of the lexicon in terms of usage.

The early loanwords from Japanese make for a significantly smaller dataset and are therefore not presented in a table. Out of the total 25 loanwords from Japanese with first attestations before 1700, 3 are monosyllabic, 15 are disyllabic and 7 are trisyllabic. There are no tokens with several syllable-variants recorded among the early loanwords. All 15 disyllabic loans are stressed on the penultimate (i.e. initial) syllable, which leaves little room for doubt about the most common stress pattern for this group. Among the trisyllabic words there is some more variation, with 3 words stressed on the antepenultimate syllable, 3 words stressed on the penultimate syllable, and 1 word with final stress. Based on this data from the early loans, it appears that my hypothesis stating that the “Italian Stress Rule” introduced by Fournier (2018) will apply to Japanese loanwords, has been supported. However, it is possible that there are differences between the early loans and the later ones, especially given the small size of this data set. Thus, final conclusions cannot yet be confidently made.

6.2 Later loans: 1701 – the present

The total number of loanwords with first attestations after the year 1701 is 1,500 tokens. The loanwords from Italian are still numerically dominant, with 999 tokens, while the Japanese loanwords amount to roughly half of that, with 501 tokens. The semantic categorisation of the later loanwords points to two particularly central areas of borrowing from Italian. Nearly 300 tokens related to music and the opera point to this as an important area of lexical import, particularly in the 1700s and the 1800s. The 1900s are clearly dominated by culinary terms, which make up nearly half the total of 288 tokens recorded in that century. This also illustrates the shift in cultural relations between Italy and the English-speaking world, as the main export from Italy shifts from cultural concepts related to music and the arts to cuisine. Italian cuisine becomes a central source of newfound interest in Italy in the 20th century, aided by the popularisation of dishes like the pizza by the Italian immigrant communities in the United States (as discussed in 4.1.2.3). *Culinary items* also represent an important semantic category for the loanwords from Japanese, with 50 tokens. Other central

semantic categories among the Japanese loanwords include *nature*, *art and material culture* and *martial arts*. A new category introduced in the latter two centuries is that of *business and commerce* which reflects the changes to the contact situation between Japan and the English-speaking world after the end of the Isolationism policy, and specifically Japan’s rise from an isolated nation to an important economic power with international influence.

In order to visualise the development in borrowing from the two source languages over time, the full distribution of tokens divided by century of first attestation in the OED is illustrated in Figure 1 below:

Figure 6.1: First attestation of loanwords by century

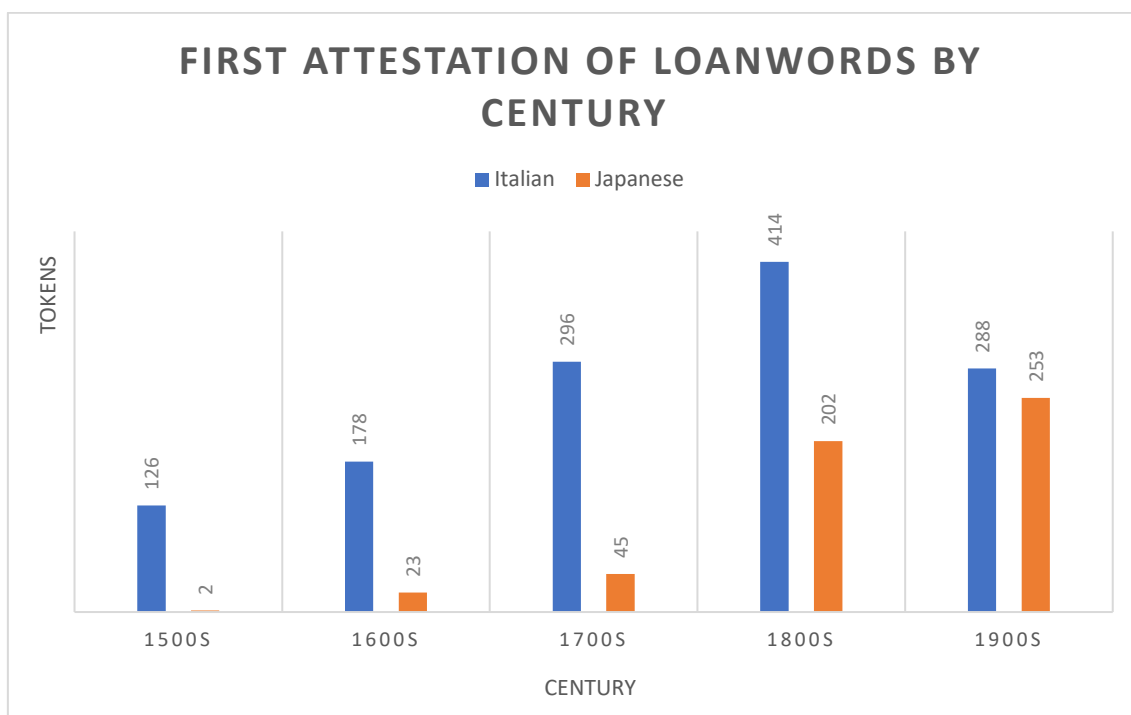


Figure 1 clearly shows the rising tendency for Japanese loanwords after the Isolationism policy ended in the 1800s. The figure also points to the 1800s as the peak century for the introduction of new loanwords from Italian, which correlates to a period in which Italian musical vocabulary was imported in great numbers. The decline observed for Italian in the 1900s can be traced back to the contact situation and loss of influence in areas such as *culture*, *arts* and *music*, which had been central semantic categories for borrowing in previous centuries. As outlined in 4.2.3, the contact situation with Japan over the last century in particular, has been one in which an increased interest in

Japanese cultural exports and food has been noticeable. It is still much too early to tell whether the trend of decline in linguistic borrowing from Italian and increase from Japanese will continue in the 21st century, or if new developments in the cultural relations with the English-speaking world will result in drastic changes in a different direction.

Another factor which reflects the change in the contact situation is the part of speech categories of the loanwords being introduced into English. While the early data from Japanese was limited to nouns only, the later period finds 16 tokens marked as adjectives as well as 5 interjections. This is indicative of a somewhat closer contact, based on the principles of a hierarchy of borrowability of categories as outlined in 2.4. The part of speech categories found among the Italian loanwords also reflects the even closer contact situation evident here, as 151 tokens are marked as adjectives³⁰, 89 as adverbs, and in addition 5 interjections, 1 conjunction and 1 preposition. Interestingly however, the conjunction, preposition and the clear majority of the adverbs all belong to the semantic category *music*. This suggests that the intensity of contact pre-supposed for the borrowing of such category items was mostly limited to that specific area of discourse.

6.2.1 Frequency

As the basis for the calculation of the OED frequency bands relies on synchronic data, differences in the frequency of specific tokens over time are not a perspective the OED can offer. However, comparing the frequency distribution of early loans with the later loanwords can still facilitate a diachronic perspective on the loanwords through comparison of the synchronic frequency of words with different first attestations in the language, and through comparing the distribution across frequency bands percentagewise.

The frequency band distribution of the loanwords with first attestations after 1701 is presented in Table 8 below. Band 7 and 8 are not included, as none of the tokens in my data were assigned to those bands.

³⁰ Note: there is overlap in the number of nouns, adjectives and adverbs, as many tokens have been listed as being used nominally and adjectivally and/or adverbially.

Table 6.4: Frequency distribution of late loans

Frequency	Italian	Japanese
<i>Band 1</i>	86	27
<i>Band 2</i>	407	193
<i>Band 3</i>	333	195
<i>Band 4</i>	122	64
<i>Band 5</i>	26	8
<i>Band 6</i>	2	-
<i>Total tokens</i> ³¹	999	501

As can be seen from Table 8 above, the total number of loanwords from Italian is almost twice as high as that from Japanese. However, Table 5 in 6.1.1. showed that the result for the early loanwords was much more numerically uneven. This indicates the strengthened position of Japanese as a source language for new additions into the English vocabulary in the later centuries.

For a different perspective on these numbers that allows for easier comparison of relative frequency distribution, Table 9 below shows the frequency distribution as a percentage, with the bands sorted into three categories as was done in 6.1.1: low bands (1 and 2), middle bands (3 and 4) and high bands (5 and 6)³².

Table 6.5: Frequency distribution of late loans: percentages

Frequency bands	Italian	Japanese
<i>Low</i>	49.3%	44%
<i>Middle</i>	45.5%	51.7%
<i>High</i>	2.8%	1.6%

Table 9 above shows that for Italian, the highest number of loanwords is to be found in the lowest two frequency bands, while the majority of Japanese loanwords belong to the two middle bands. The numbers above place words of Italian origin slightly ahead of those from Japanese for the percentage of the total set of loanwords which are found in higher frequency of use. However, a comparison with the early period shows that the

³¹ The total number includes 23 tokens from Italian and 14 tokens from Japanese where no frequency band was listed.

³² The tokens without frequency band recorded (N/A) are not included in Table 9, but amount to 2.3% for Italian and 2.8% for Japanese.

numbers have levelled significantly, with only a 1.2 percentage point difference, compared to the 10.3 percentage point difference for the early loanwords. This indicates that the borrowing situation has become more balanced in terms of the frequency in textual use among the total tokens across the two source languages.

6.2.2 *Phonological adaptations*

As seen among the early loans, the strong trend of vowel breaking of unstressed /o/ in final open syllables continues to be dominating in the later loans from both source languages. A total of 388 tokens from Italian end in <o>, which according to the conventions of Italian grapheme to phoneme correspondence would be pronounced as /o/. Of these, a clear majority totalling 340 tokens display vowel breaking in English and are transcribed with either /əʊ/ (RP) or /oʊ/ (GA). There are 53 outliers that retain /o/ in their listed transcription, and among these the majority belong to the lower frequency bands 2 and 3. Another feature that characterises these words is their broad semantic categorisation, as 27 tokens belong to the semantic category *music*. The fact that they belong to a specialised discourse, that of musical jargon, points to the possibility that specialised technical language is placed in the periphery of the lexicon generally.

The number of tokens with word-final <o> among loanwords of Japanese origin amounted to 77, where 70 displayed the trend of vowel breaking to either /əʊ/ (RP) or /oʊ/ (GA). The 7 tokens that show resistance to this trend and are transcribed with /o/ are divided between frequency bands 2, 3 and 4. No single semantic category stands out for this group, but they are all related to aspects of Japanese nature, society or material culture. The fact that the referents for these terms are culturally specific concepts could perhaps be seen as a contributing factor or explanation of their status as phonologically unassimilated. However, as the referents of the majority of all tokens of Japanese origin³³ can also be described as more or less ‘culturally specific’ this can hardly be considered as a determining factor.

³³ Culinary terms and martial arts are some examples of categories not necessarily ‘culturally specific’ to Japan at this point in time, as the referents have also been exported. The concept of ‘cultural specificity’ is fairly vague in the globalised reality of today though.

The trend of reduction of unstressed word-final <a> (/a/³⁴) to schwa also continues to be evident in the data from the later period. A total of 87 tokens of Japanese origin end with <a>, and of these as many as 72 follow the trend of reduction to /ə/. The 15 words that are listed with /a/ do not appear to have any particular semantic characteristics, and are divided between band 2, 3 and 4. For Italian, 303 words ended in <a> and among these, 275 were listed with pronunciation variants which included reduction to schwa /ə/. (how many maintained a?)

A further detailed look at this group finds that 43 of the tokens ending in <a> in fact end in the sequence <ia>, which can be considered a diphthong in Italian (see 5.1.1.1 for the discussion of Italian diphthong formations). The question of how English speakers would interpret these constructions was posed in 5.1.1.1 as well, as English does not have /ia/ in its inventory of diphthongs. The results show some variation among the possible adaptations of this sequence, but the main strategies of adaptation of <ia> are the centring diphthongs /iə/ or /ɪə/, observed in 21 tokens, and the division into two syllables by producing a long vowel followed by schwa /i:ə/, observed in 11 tokens. A third adaptation strategy exhibited by 7 tokens reflected an interpretation of the <ia> sequence as a glide + vowel construction, such as in *Rosalia* /roʊ'zəlɹjə/ (GA variant). However, for some of the tokens in this group, the use of /j/ is in part conditioned by the preceding grapheme <gl>.

The grapheme <gl>, which in Italian is used to represent the lateral approximant /ʎ/ that is not found in English, is present in 12 tokens from the Italian data. The most common adaptation strategy for this phoneme, is a lateral /l/ followed by an approximant /j/, as in *scaglia* /'skɑ:lɹjə/. Another adaptation strategy, found in variants listed for RP, involves a lateral /l/ followed by the centring diphthong /ɪə/, such as in *tagliatelle* /taglɪə'tɛli/.

The word *tagliatelle* /taglɪə'tɛli/ also illustrates adaptation of word-final <e> to /i/ in pronunciation. As seen for the early data, the adaptation of <e> was not as regular as the other final vowels, and four different strategies were evident. For the late Italian loan words, the same four strategies were found, and among the 185 tokens ending in <e>,

³⁴ likely to be produced as /a/ in the source languages given their vowel inventories and regular relationship between grapheme and phoneme.

62 included variants where <e> was pronounced as monophthong /i/, while 77 tokens included variants with vowel breaking to /ei/. 11 tokens reproduced <e> as /ɛ/. The remaining occurrences of <e> either involved apocope or were silent due to other factors. For the Japanese loanwords, vowel breaking and /i/ were also the most common strategies, with 22 tokens including variants with /i/ and 17 including /ei/. 5 tokens were attested with /ɛ/, and a mere 3 tokens included a silent final <e>.

As seen in the early data from Italian, the concept of consonant gemination is again recorded among the loanword from the late period. The majority of the tokens displaying this feature in the OED transcription belong to the semantic category of *music*, and also notably resist other regular processes of nativization such as vowel breaking and the reduction of final /a/ to schwa. The presence of tokens that display illicit features such as gemination in middle frequency bands, such as *basso* /'basso/ in band 4, problematises the idea of the frequency bands as related to the level of nativization, as the expectation would be to find these words in the lowest bands.

The liquid phoneme in Japanese shows consistent mapping to the rhotic /r/ in English, with only two appearances of the lateral /l/, both in words that have been combined with an English element: *urushiol* 'An oily liquid causing skin irritation, obtained from the Japanese lacquer tree' (Etymon: *urushi* + English combination form *ol*), and *andosol* 'A mineral-rich soil derived from volcanic matter' (etymon: *ando* + English combination form *sol*). These are both scientific terms, from the area of chemistry and soil science respectively, and contain combination forms from these fields of discourse, which means the /l/ is not attested in elements based on a Japanese form. The idea suggested in 5.2.1.3, that English speakers might interpret the Japanese acoustic signal as more similar to lateral /l/ than the rhotic thus appears wholly unsubstantiated. This could also point to the fact that English speakers are affected by the orthographic representation of phonemes, as the liquid phoneme in Japanese is consistently transcribed using <r> in the most common romanisation systems and is only represented as <l> in the two exceptions listed above.

6.2.3 Orthographic adaptations

Many of the same regular processes of orthographic adaptation observed among the early loanwords in 6.1.3 are also found in the data from the later period, such as the loss

of final vowel graphemes to match apocope of the final vowels from the source language etyma, as in the Italian loanword *Romagnol* (RP: /,rɔmənə'njəʊl/) from the etymon *romagnolo*.

A noticeable trend across both Italian and a few Japanese loanwords concerns the pronunciation of the word-final sequence <ite> as /ɪt/ (RP) or /aɪt/ (GA). On the surface, it could appear to be a case of orthographic effect through analogy with native items, but as seen in the two examples containing /l/ above, the semantic field plays a crucial role. A total of 14 tokens from Italian and 2 from Japanese contain this sequence, and they all belong to the semantic category of *mineralogy*, in which the *-ite* suffix is used as a combination form to coin new words for types of minerals. It is thus not a reflection of a change of a source language form, but a productive process involving a suffix used universally in a specific discourse, which is routinely pronounced as /ɪt/ (RP) or /aɪt/ (GA).

As discussed in 6.1.3, orthographic effect is often a result of adaptation through analogy with other orthographically similar forms in the L1. Another example of this process is evident in Italian loanword *pizzelle* ‘A type of fried, often stuffed, dumpling, also: small pizza’, listed with the pronunciation variants /pi:'tseɪl/ and /pi:'tsɛl/. As suggested by the final <e>, the form is based on the plural of the original Italian etymon *pizzella*. Through analogy with other loanwords in English with the French suffix *-elle*, originally from French, the final syllable has been elided in the second variant /pi:'tsɛl/.

6.2.4 Morphological adaptations

As observed in 6.1.4, the number of tokens with inflectional information specified represents a minority of the total data. Among the 999 Italian tokens from the late period, plural inflectional information was specified for 256 tokens, while 118 of the 501 tokens of Japanese origin included this information. While the majority of the total tokens are nouns³⁵, the 1126 words across both source languages with no plural marking specified also include words from other part of speech categories. 90 of the tokens originating in Italian and 34 tokens from Japanese are divided among adverbs and adjectives, as well as few verbs, interjections and prepositions.

³⁵ Either only nouns or with noun as one of the listed categories. Some adjectives are used nominally etc.

Based on the observations of the early data in 6.1.4, the following three scenarios were introduced regarding plural marking: tokens retain source language plural, tokens show variation between source and target language plural, and finally the tokens take only the target language plural *-(e)s*. Based on those three scenarios, a cline of nativization was proposed, but evidence of regular plurals being attested in early variants and subsequently replaced through a re-introduction of source language plural marking posed a problem for the assumption of this cline as a unidirectional development from foreign to native.

Among the 256 tokens from Italian with plural marking specified, only 5 tokens are listed with English *-(e)s* plural marking only. However, the loanwords listed with only Italian *-i* or *-e* plurals make up a significant portion of the total with 104 tokens. Despite the clear numerical difference, it is not easy to define any special characteristics of either group. A correlation between the frequency bands and the choice of plural marking cannot be adequately determined, as the tokens in the group retaining source language plurals ranges from band 1 to 5, while the few tokens in the target language plural group are spread among band 2, 5 and 6. The single token in band 6 is hardly enough to make a claim that higher frequency relates to a higher chance of taking target language plural marking.

In the group listed with several variants for plural marking, further detailed distinction is possible. Both unchanged and Italian *-i* or *-e* plurals were attested for 3 tokens, while 6 tokens had both unchanged and English *-(e)s* plurals listed. In the latter group we find several tokens ending in *<i>*, which indicate plural form in the source language. The variant including plural *-s* thus results in an irregular form with double plural marking, such as the variant *cannellonis*. This indicates that English speakers are not necessarily always sensitive to the plural marking of the source language in a borrowing situation. Another irregularity is found in the word *braccio*, which is listed with *braccia* as the plural form. This construction is neither a reflection of a regular plural formation strategy in Italian or in English³⁶, and illustrates the fact that irregularities sometimes occur in transmission of linguistic material.

³⁶ Arguably “regular” in English through Latin loanwords ending in *-um*, such as *stratum – strata* etc, which could be transferred to other lexical items through analogy.

The largest group, consisting of 113 tokens, had both source language and target language plurals attested. As seen in the source language plural group, the frequency bands range from 1 to 5, and so a clear correlation is again difficult to ascertain. Additionally, 23 tokens were specified as taking plural agreement, and of these the majority were culinary terms used as mass or non-count nouns and ending in <i>, such as *ravioli*, *crostini*, *pepperoni* and *salami*. As evidenced by the final <i>, these words have been borrowed in their plural form from Italian and have kept that meaning and form in English.

If we go back to our proposed cline of nativization, the numbers observed for the late loanwords would suggest that very few words have reached the ‘core’ of the lexicon in terms of regular English plural -s marking. However, highly frequent items tend to retain morphological irregularity. The inflection of the verb *be* is a prime example of this, as is the irregular plural formation evident in the noun *child* – *children*, which belong to frequency bands 8 and 7 respectively and are undoubtedly part of the core area of the English lexicon. This complicates the picture presented in the proposed cline of morphological nativization further, as regularisation of morphology is apparently not necessarily tied to high frequency or specific strata of the lexicon.

Plural marking on early loanwords from Japanese was not recorded on a sufficient number of tokens to identify any particular trend, but for the later loans, some patterns can be noted. For 45 tokens, the plural form is specified as being unchanged, which reflects the lack of morphological plural marking in the source language. As observed for the corresponding group of source language plurals among the Italian loanwords, the frequency bands range from 1 to 5, and so a definitive correlation between these factors is not traceable here. 4 other tokens have been specified as taking plural agreement either usually or occasionally in their plain form.

Interestingly, there were a mere two tokens listed with English -(e)s plural only. One of these was the word *honcho*, classified by the OED as colloquial language, which was brought into use by American prisoners of war in Japan during WWII. The unique etymology and its status as colloquial might be factors contributing to its regular -s plural in this case. Furthermore, among the 66 tokens listed with both unchanged and -(e)s plurals as attested variants, several specify that English -(e)s plural is rare or only

occasionally used, such as for *minshuku*, *maiko* and *samurai*. This could suggest that English speakers are sensitive to the fact that Japanese does not tend to mark plurality morphologically, and that this sensitivity and desire to faithfully reproduce a source language form competes with nativization processes. If we go back to the cline outlined in 6.1.4, the observations above would lead to a conclusion that none of the loanwords from Japanese have moved all the way towards the core stratum of the lexicon in terms of inflectional morphology, as none of them take the regular L1 native plural form only. However, it is again important to keep in mind that there are highly frequent native vocabulary items with irregular morphological plural marking, and unmarked plurals also exist in native ‘core’ vocabulary – plural marking alone can thus not serve as the only measure of placement in a core-periphery model.

6.2.5 Syllable distribution and stress pattern

Among the total 999 loanwords from Italian introduced after 1701, 5 are monosyllabic, 137 are disyllabic, 361 are trisyllabic, 328 are tetrasyllabic and 140 are polysyllabic, i.e. with 5 or more syllables. Additionally, 28 words have a variable number of syllables recorded in the different pronunciation variants listed in the OED. The stress pattern data for the later loanwords from Italian distributed by number of syllables is presented in Table 10 below:

Table 6.6: Stress placement on late Italian loans

	Disyllabic	Trisyllabic	Tetrasyllabic	Polysyllabic	Variants
<i>Final</i>	10	12	1	1	1
<i>Penult</i>	126	286	262	82	12
<i>Antepenult</i>		56	42	21	-
<i>Preantepenult</i>			8	2	-
<i>Variant</i>	1	7	16	34	14

The numbers in Table 10 above show a clear preference for placing stress at the penultimate syllable for Italian loanwords across all syllabic groups. For the trisyllabic, tetrasyllabic and polysyllabic groups, a total number of 119 tokens were listed with antepenult stress. A common feature among them is a light penult syllable, which is consistent with the strong trend observed by Krämer (2009) that heavy penultimate

syllables³⁷ attract stress (as discussed in 5.1.2). Only when no such heavy syllable is found, the stress might fall on the antepenult syllable. This confirms the trend observed in the early data. A notable factor in common for the words with final stress, in particular the disyllabic and trisyllabic words, totalling 22 tokens, is a closed final syllable. As discussed in 5.1.2, Italian strongly prefers open final syllables, and the presence of closed ones appears to be the result of a combination of adaptations such as apocope or orthographic effect through analogy with suffixes in L1. One example of apocope being evident in *pastorale*: RP variant /,pastə'ra:l/.

The variable stress group includes a number of tokens that are made up of several words, such as *primo tenore assoluto*, GA: /' ,primou tə'nɔri ,asə'l(j)udou/, RP: / ,pri:məu tɛ ,nɔ:ri asə'l(j)u:təu/, and *pizza alla Napoletana*, RP: /' pi:tsə(r) əl ə nə ,pɔlɪ'tɑ:nə/. It is hard to conclusively say if English speakers routinely interpret such sequences as phrases or as long compounds in terms of stress placement. The OED lists a variation of primary stress placements for these constructions, as can be seen in their transcription variants, where the RP variant of *primo tenore assoluto* has primary penult stress on the final word, while the GA variant has primary stress placed on every word.

The 27 tokens listed with several variants for syllable count reflect differences in the placement of syllable boundaries based on re-interpretation of certain vowel sequences as diphthongs or glide + vowel segments versus as sequences forming separate syllable-nuclei (as discussed in 5.1.1.1 and 6.2.2.). One such example of variation in syllable count is found in the versions listed for the Italian loanword *soave*: /səu'ɑ:veɪ/ (trisyllabic) and /'swɑ:veɪ/ (disyllabic).

The data for the early period of Japanese loanwords discussed in 6.1.5 was scarce, and the stress assignment showed a strong trend of stress placement on the penultimate syllable, which was consistent with the hypothesis of the “Italian rule” for stress placement on foreign loanwords. The stress placement on the significantly larger group of later Japanese loanwords is presented in Table 11 below:

³⁷ Heavy syllables end in a consonant or diphthong while light syllables end in a monophthong.

Table 6.7: Stress placement on late Japanese loans

	Disyllabic	Trisyllabic	Tetrasyllabic	Polysyllabic	Variants
<i>Final</i>	9	3	2	-	-
<i>Penult</i>	183	82	72	20	3
<i>Antepenult</i>		52	2	1	-
<i>Preantepenult</i>			2	1	-
<i>variant</i>	10	15	7	2	3
<i>No info</i>	-	2	1	-	-

As expected, based on the results for the data in the early section, the numerically dominant stress pattern is penultimate stress for every syllabic category. For the trisyllabic group, antepenult stress was also a notable trend, illustrating that for words of this syllabic number, initial stress is also a productive adaptation for imported words with no stress pattern in the source language. The total 37 tokens with a variable stress pattern include words where variants of pronunciation included penultimate as well as either final or antepenult stress. Whether there is a diachronic aspect to stress assignment in variants where one pattern has been introduced later or represents a later, more regular process than another is unfortunately not possible to elicit from the OED data alone.

The three total tokens with no stress placement information indicated in the OED all belong to the middle to lower frequency bands (band 2 and 3 respectively), which could contribute to the indication that these words belong to the periphery of the lexicon, as they maintain the source language feature of no stress. However, it is also possible that these entries are simply the result of information being left out in the OED's records for other arbitrary reasons.

6.3 Discussion of linguistic analysis

The main aim of my thesis, as formulated in my first research question (R1), has been to try to uncover trends in the nativization of loanwords from the different source languages Italian and Japanese into English. The results presented in 6.1 and 6.2 makes it clear that there are indeed trends in the nativization of loanwords, as predicted by my first hypothesis (H1). The fact that these trends to a significant degree base themselves

on rules governing the L1 phonological system as well as orthographic convention further confirms H1. Some of the main trends in phonological adaptation include vowel breaking of unstressed /o/ to /əʊ/ for RP and /oo/ for GA, reduction of unstressed word-final /a/ to schwa /ə/ and alternation between realisation of word-final <e> as /i/ or /ei/. Adaptations through orthography frequently involved analogy with suffixes in the L1, such as the sequence <ese>, pronounced as /i:z/.

The preliminary conclusion formed in 4.3 regarding the validity of H2, was not confirmed through the linguistic analysis. Evidence of influence from the cultural contact situation in the processes of linguistic nativization was not found outside of a few special cases. As noted for H1 above, the linguistic processes involved in nativization were largely the same for both source languages. The differences in the contact situation primarily manifested themselves through the differences in the scope of borrowing, through the number of tokens to the part of speech categories found among the data. The presence of a considerable number of adverbs related to the semantic field of *music* in the late Italian data is particularly indicative of a more intense influence from the cultural contact in that area of discourse.

Another aspect of my project has been an analysis of the impact of differences in source language typology on nativization. The question of whether adaptation processes would be sensitive to typological differences was posed, as formulated in RQ 3. My prediction in H3 was that typological differences would prove to be less important than the L1 system, and that similar adaptation processes would thus apply to loanwords from both source languages. The presence of the same main phonological trends such as vowel breaking, reduction or loss in loanwords from both languages is a factor which appears to satisfactorily confirm this hypothesis. An element of sensitivity to source language typology was however traceable morphologically in plural marking, where the most frequent strategies for the later period involved variation between source language plural and English -s or faithful reproduction of plural marking patterns found in the source language, i.e. -i or -e for Italian loanwords and unchanged plurals for loanwords of Japanese origin. The same processes of approximation, mechanic adaptation and analogy involving phonology and orthography all appear to affect loanwords from the two languages to an equal degree. There were features of the source language typologies observable in the outliers that had resisted these processes, but the presence of these

exceptions still does not refute the hypothesis that the L1 system is the most important factor in nativization, and that the general trends do not show sensitivity to typological difference.

RQ 4 concerned the use of the OED frequency bands to relate lexical items to a relative placement in the core-periphery organisation of the lexicon. One of the limitations to using the frequency bands in this way is firstly the fact that the correlations between linguistic features associated with the ‘core’ of the lexicon and frequency are not always clear, as seen in the analysis of plural marking in 6.1.4 and 6.2.4. Another limitation is the fact that the frequency bands themselves are not reflective of the “real” frequency of a word but are limited to *textual* frequency. This leaves out the use of words in more informal registers of language such as online communication or oral discourse. A third limitation is the fact that the data underlying the calculations of the bands is from 2012, which means the frequency of use in the last 7 years is not reflected. A prime example of a word whose frequency band is a clear mismatch with a more realistic frequency of use is *emoji* (OED: 1997), which is classified as band 1. The description of the words in band 1 as “highly obscure” hardly seems applicable to this word in its current use. The phenomenon of vowel breaking observed in the pronunciation, as well as the option of English plural *-s* marking further serve as indicators that this word has moved further along the path of nativization and belongs closer to the core of the lexicon than its band 1 status would suggest. Based on these limitations, I would argue that H4, stating that these bands could be employed as a useful measure of the relative positioning of lexical items along a core-periphery organisation of the lexicon cannot be confirmed, as I did not find a consistently strong correlation between the levels of nativization and the frequency band.

The hypothesis (H5) posed for RQ 5, that stress will be placed on the penultimate syllable for Japanese loanwords ending in an open syllable following the Italian Stress Rule appears to have been strongly substantiated by the evidence presented in both 6.1.5 and 6.2.5. However, a comparatively high number of trisyllabic tokens displaying antepenult stress, presents this a common alternative stress pattern for words of this syllabic group. The stress patterns of Italian loanwords also predictably favoured the penultimate stress pattern, as expected based on the findings of previous research by Fournier (2018).

The hypothesis formulated in H6, that early loanwords will show a higher degree of nativization, and exhibit fewer markedly foreign features than later loanwords, appears through the diachronic perspectives offered by my linguistic analysis to be only partly confirmed. There does not appear to be a clear difference in the nativization of early loans versus later loans to be observable in the synchronic data in terms of phonological or orthographic processes. The main indication of a difference that could potentially be due to time, is found in the data with inflectional information specified. The presence of 28 tokens with English *-(e)s* plural marking only in the early Italian data compared to a mere 5 tokens for the later data indicates that English inflection is more common for loanwords which have been in use in English for longer. However, the diachronic perspective offered by information on historical variant spellings indicated that the process of assigning English inflection was regular to a somewhat higher degree in the early period, and that later variants re-introduced foreign plural marking. Due to the limitations of this later trend of re-introduction of foreign plurals to a small number of tokens, I nevertheless conclude that H6 is partly substantiated, in that there are certain limited indicators of time of admission playing a role in nativization, specifically in the realm of plural marking.

7 Conclusion

The present thesis has analysed and discussed some of the trends in the nativization of Italian and Japanese loanwords into English as recorded in the OED. Through including factors such as frequency, typological differences between the source language and the borrowing language as well as the sociocultural context of the language contact in the analysis, new perspectives have been offered on the linguistic adaptations involved in nativization. By comparing trends in the nativization of linguistic material in the areas of phonology, orthography, morphological plural marking and stress assignment for loanwords from two different source languages, my study shows that the processes involved are not unique to a single borrowing situation.

The outline of the typological features of the source languages illuminated the areas in which linguistic adaptation would be needed in order for words to conform to the English systems. The overview of the linguistic features of the source languages further served to highlight areas where re-interpretation and analysis of material could lead to discrepancies between borrowings and their source language etyma. One such discrepancy involved the re-analysis of sound sequences into segments leading to resyllabification, which in some cases subsequently affected stress placement. While some of the adaptations were necessitated by the specific typological features that marked words as foreign, such as the different specific foreign phonemes, the adaptation processes involved in changing these to acceptable forms in L1 were applied across source languages with no particular sensitivity to underlying typological features of the source languages. Following the observations of previous studies, I found that the linguistic adaptations could be attributed to the same underlying processes, such as approximation, mechanic adaptation and analogy. The trends found in the phonological adaptations were particularly regular, as a clear majority of tokens from both source languages underwent the same processes of vowel breaking of unstressed /o/ and reduction of /a/ to /ə/.

Orthographic effects were expected to be found to a greater extent among loanwords from Japanese, especially for the early loans, given the geographical distance to the English-speaking world, which put natural limits on early direct contact. This

limited direct contact would presumably subsequently lead to an increased reliance on orthographic input in the absence of exposure to the acoustic signal from the source language. However, the trends observed in the phonological adaptations of Japanese loanwords did not show any particular increased orthographic effect beyond that observed for Italian. The spelling of Japanese loanwords in the synchronic data further appears to have been largely based on standards for romanisation of Japanese, and do not show considerable differences from the transliterated versions of their source language etyma. The spelling of Italian loanwords was also regular in the majority of cases, and the main trend was clearly to reproduce the orthographic form from the source language faithfully. One limited type of orthographic adaptation included loss of word-final graphemes to reflect apocope, such as in *arsenal* RP: /'ɑ:sən(ə)l/ from the Italian etymon *arsenale*. Some orthographic effects were noticeable where grapheme combinations in the source languages were re-interpreted through analogy as native suffixes, such as in words ending in <ese>, reproduced as /i:z/.

The thesis set out to test the specific hypothesis of “Italian stress” on the penult syllable as a pattern used for foreign stress, through comparing Italian stress with the stress assignment on Japanese loanwords. Where previous studies were limited to words with a specific phonotactic structure, my analysis included all tokens in order to see trends in the stress patterns across the full dataset. Evidence from the 526 tokens from Japanese showed that a clear majority across all the syllabic categories assigned stress on the penultimate syllable, thus confirming the hypothesis of the “Italian rule” for loanword stress assignment in words from Japanese.

Frequency was a factor my thesis aimed to include in order to facilitate a new perspective on these loanwords compared to the previous studies. However, I found that while the frequency information provided by the OED frequency bands was useful in comparing the relative use of loanwords between the two source languages, the bands had some clear limitations in relation to further theories. The correlation assumed between the core-periphery theory and a cline of high to low frequency of use, as indicated by the OED frequency bands, proved to be difficult to ascertain. The loanwords that exhibited resistance to the processes of nativization through preserving features from the source language that violated the constraints active on native L1 vocabulary were not necessarily consistently associated with a low frequency band.

The idea that the nativization of loanwords can be likened to a path or cline where words move from unassimilated foreign vocabulary in the periphery towards a more assimilated “core” of the vocabulary where words are no longer perceived as foreign was used as a basis for the proposition of a cline for the marking of plural inflection. However, this cline is ultimately a simplification of the nativization process and does not specify the various factors that might motivate speakers to maintain foreign features. It also suggests that the process of nativization moves in one direction, from foreign to assimilated, which is not necessarily the case. This fact that can be seen through the ‘correction’ of plural marking that appears to have happened over time for some tokens, where loanwords have been changed to appear *more* markedly foreign after having been nativized to assimilate with native elements. The evidence of the loss of regular English -s plural inflection on early Italian and Japanese loans over time, as discussed in 6.1.3 thus appears to tell a story of a more winding path of nativization than what the presentation of this process as a “cline” suggests.

Although difficult to measure, the social factor is arguably of importance here as well and motivates language users to maintain faithfulness to the original form in order to maintain clear ties to the etymology of the loanword and activate associations to the source language culture. The use of Italian terms was, and arguably still is, connected with a certain element of prestige, and their use extends beyond just filling a lexical gap. It is also a fact that a large number of loanwords belong to specific areas of discourse, such as technical vocabulary related to music from Italian, and terms used in the practice of martial arts from Japanese. The affiliation with specific discourses can also be a factor which motivates speakers to rank faithfulness to the source language form higher than other constraints on well-formedness and markedness in their L1.

While the extent and intensity of linguistic contact between English and the two source languages would be classified as casual in both cases, it is nevertheless clear that the cultural and linguistic impact left by Italian has been considerably deeper. The higher total number of tokens imported as well as the wider variety in parts of speech categories of the borrowings from Italian compared to Japanese are both results of a deeper cultural contact across various fields of discourse, and particularly in the areas of art, music and cuisine.

However, the contact with and interest in Japan has increased over the past century, which an increasing number of loanwords imported from Japanese in the 1900s reflects, specifically in the fields of cuisine and martial arts, and more recently digital culture. Judging by this tendency, it appears entirely plausible that Japanese popular culture can find a broader audience in the English-speaking world. This could, perhaps helped along by a more conscious effort on the part of the Japanese government in promoting and spreading the language, lead to an even greater influx of Japanese loans and an impact that may rival that of Italian in time.

7.1 Avenues for further study

A potentially interesting avenue for further study would be testing the findings of a dictionary-based study such as my own, with data collected through other means, either through language corpora of different types or through oral or written tests of respondents through surveys or in an experimental setting.

While some studies have been conducted on areas of loanword integration and adaptation through an experimental framework with test subjects, such as studies on stress placement and linked to orthographic effect, there are also other aspects of loanword nativization that could be interesting to study further in such ways. Studies of the nativization of loanwords in an experimental setting including speakers with different linguistic backgrounds and levels of bilingualism could add more detailed perspectives on variations in pronunciation, stress placement, inflection and more. As a large number of tokens in my dictionary-based data showed variation in plural marking, a study of actual speakers and their choice of inflection on such loanwords attested as variable could perhaps help elucidate the issue further, by uncovering possible further trends in inter-speaker variation.

My study found the information on frequency offered through the OED frequency bands to be of limited usefulness for my theories, but the use of frequency measures in language corpora could add an interesting dimension to studies of loanwords in use. A comparison between corpora drawing their data from different fields of discourse or from different varieties of English could facilitate a comparison between the integration and frequency of use of loanwords in specialised fields versus

in general discourse and across varieties of English globally. This could also add another perspective to discussions surrounding the differentiation between referential use of foreign language or codeswitches and borrowing.

References

- Bertinetto, Pier Marco, and Michele Loporcaro. 2005. 'The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome', *Journal of the International Phonetic Association*, 35: 131-51.
- Calabrese, Andrea, and Leo Wetzels. 2009. *Loan phonology* (John Benjamins: Amsterdam).
- Cavaioli, Frank J. 2008. 'Patterns of Italian Immigration to the United States', *Catholic Social Science Review*, 13: 213-29.
- Chaney, Edward. 1998. *The Evolution of the Grand Tour : Anglo-Italian Cultural Relations since the Renaissance* (London: Routledge: London).
- Clulow, Adam. 2013. 'Commemorating Failure: The Four Hundredth Anniversary of England's Trading Outpost in Japan', *Monumenta Nipponica*, 68: 207-31.
- Coetsem, Frans van. 1988. *Loan phonology and the two transfer types in language contact* (Foris: Dordrecht).
- Cossu, Guiseppe. 1999. 'The acquisition of Italian orthography.' in Margaret Harris, Giyoo Hatano, George Butterworth and Kurt W. Fischer (eds.), *Learning to Read and Write: A Cross-Linguistic Perspective* (Cambridge University Press: Cambridge).
- Delahunty, Andrew. 2008. *From bonbon to cha-cha : the Oxford dictionary of foreign words and phrases* (Oxford University Press: Oxford).
- Durkin, Philip. 2014. *Borrowed words : a history of loanwords in English* (Oxford University Pres: Oxford).
- Duus, Peter. 1998. *Modern Japan* (Houghton Mifflin: Boston).
- Fournier, Pierre. 2018. "Stress Assignment in Italian Loanwords in English and its Impact on the Stressing of Foreign Words by Native English Speakers." In *Corela*. Cercle linguistique du Centre et de l'Ouest - CerLICO.
- Friesner, Michael L. 2009. 'The adaptation of Romanian loanwords from Turkish and French.' in Andrea Calabrese and Leo Wetzels (eds.), *Loan phonology* (John Benjamins Pub. Co.: Philadelphia, PA).
- Gordon, Andrew. 2003. *A modern history of Japan : from Tokugawa times to the present* (Oxford University Press: New York).
- Groot, Gerry. 2018. 'Cool Japan Versus the China Threat: Does Japan's Popular Culture Success Mean More Soft Power?' in, *Japanese Language and Soft Power in Asia* (Springer).
- Gusmani, Roberto. 1973. *Aspetti del Prestito Linguistico* (Libreria Scientifica Editrice Napoli).
- . 1987. 'Interlinguistica.' in R. Lazzeroni (ed.), *Linguistica Storica* (Nuova Italia Scientifica Roma).
- Haberland, Detlef. 2012. "Kaempfer, Engelbert." In *Encyclopædia Iranica Online*. Encyclopædia Iranica Foundation, INC. .
- Hanlon, Gregory. 2000. *Early modern Italy, 1550-1800 : three seasons in European history* (Macmillan: Basingstoke).
- Harrington, Jonathan, and Felicity Cox. 2019. "The syllable and phonotactic constraints." In *Phonetics and phonology resources*. Macquarie University Macquarie University: Department of Linguistics

- Hartmann, Reinhard Rudolf Karl, and Francis Colin Stork. 1972. *Dictionary of language and linguistics* (Applied Science Publishers LTD: London).
- Hashimoto, Kayoko. 2018. 'Cool Japan and Japanese Language: Why Does Japan Need "Japan Fans"?' in, *Japanese Language and Soft Power in Asia* (Springer).
- Haugen, Einar. 1950. 'The Analysis of Linguistic Borrowing', *Language*, 26: 210-31.
- Helstotsky, Carol. 2008. *Pizza : a global history* (Reaktion Books: London).
- Hunt, Jocelyn. 1999. *The Renaissance* (Routledge: London ; New York).
- Istat. 2018. 'Movimento Turistico in Italia ', Istituto Nazionale di Statistica, Accessed February 2019. <https://www.istat.it/it/archivio/224376>
- Itô, Junko, and Armin Mester. 1995. 'The Core-Periphery Structure of the Lexicon and Constraints on Reranking', *University of Massachusetts Occasional Papers*, 18.
- Kaislaniemi, Samuli. 2017. 'The Linguistic World of the Early English East India Company: A Study of the English Factory in Japan, 1613–1623', *Journal for Early Modern Cultural Studies*, 17: 59-82.
- Karan, Pradyumna. 2005. *Japan in the 21st century : environment, economy, and society = 21 seiki no Nihon* (University Press of Kentucky: Lexington).
- Krämer, Martin. 2009. *The phonology of Italian* (Oxford University Press: Oxford).
- Kubozono, Haruo (ed.)^(eds.). 2015. *Handbook of Japanese phonetics and phonology* (De Gruyter Mouton: Boston).
- Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages* (Blackwell: Oxford).
- Lawrence, Jason. 2006. *'Who the Devil Taught Thee So Much Italian?' : Italian Language Learning and Literary Imitation in Early Modern England* (Manchester University Press: Manchester).
- Maiden, Martin, and Cecilia Robustelli. 2007. *A reference grammar of modern Italian* (Hodder Education: London).
- Matras, Yaron. 2007. "The borrowability of structural categories." In *Grammatical borrowing in cross-linguistic perspective*, edited by Yaron Matras and J Sakel. ———. 2009. *Language contact* (Cambridge University Press: Cambridge).
- Michel, Wolfgang 2005. 'On Engelbert Kaempfer's "Ginkgo": 1-5.
- Ohata, Kota. 2004. 'Phonological Differences between Japanese and English: Several Potentially Problematic Areas of Pronunciation for Japanese ESL/EFL Learners', *Asian EFL Journal*, 22: 29-41.
- Otake, Takashi. 2015. "Mora and mora-timing." In *Handbook of Japanese phonetics and phonology*, edited by Haruo Kubozono. Boston: De Gruyter Mouton.
- Pinnavaia, Laura. 2001. *The Italian Borrowings in the Oxford English Dictionary: A lexicographical, linguistic and cultural analysis* (Bulzoni Editore: Rome).
- Sansom, G. B. 1978. *A history of Japan : 3 : 1615-1867* (Dawson: Folkestone).
- Serjeantson, Mary S. 1935. *A history of foreign words in English* (Kegan Paul, Trench, Trubner: London).
- Shibatani, Masayoshi. 2017. "Japanese language." In *Encyclopaedia Britannica*. Encyclopædia Britannica, inc.
- Shimizu, Katsumasa, and Masatake Dantsuji. 1987. 'A. Cross-Language Study on the Perception of [r-l]: A Preliminary Report', *Studia phonologica*, 21: 10-19.
- Simone, Maria Rosa di. 2003. 'Students.' in Hilde de Ridder-Symoens and Walter Rüegg (eds.), *A History of the University in Europe: Volume 2, Universities in Early Modern Europe (1500-1800)* (Cambridge University Press: Cambridge).

- Svensson, Ann-Marie. 2004. 'On the stressing of French loanwords in English.' in Christian Kay, Carole Hough and Irené Wotherspoon (eds.), *New Perspectives on English Historical Linguistics II*.
- Sweet, Rosemary. 2012. "Cities and the grand tour : the British in Italy, c.1690-1820." In. Cambridge ; New York: Cambridge University Press.
- Thomason, Sarah Grey, and Terrence Kaufman. 1992. *Language contact, creolization, and genetic linguistics* (University of California Press: Berkeley).
- Tsujimura, Natsuko. 2007. *An introduction to Japanese linguistics* (Blackwell: Malden, Mass).
- . 2013. *Introduction to Japanese Linguistics* (Wiley: Somerset).
- UNESCO. 'Country page: Italy', Accessed February 2019. <https://whc.unesco.org/en/statesparties/it>
- University of Tokyo. 2009. 'Recommended System for Romanizing Japanese', The University of Tokyo Accessed April 2019. <http://park.itc.u-tokyo.ac.jp/eigo/UT-Komaba-Romanization-of-Japanese-v1.pdf>
- UNWTO. 2018. 'UNWTO Tourism Highlights 2018 Edition', UN World Tourism Organisation, Accessed April 2019. <<https://www.e-unwto.org/doi/pdf/10.18111/9789284419876>>
- Vendelin, Inga, and Sharon Peperkamp. 2006. "The influence of orthography on loanword adaptations." In *Lingua*, 996-1007.
- Wells, J. C. 2008. *Longman pronunciation dictionary* (Pearson/Longman: Harlow).

Dictionaries:

English:

'Oxford English Dictionary Online'. Oxford University Press. Accessed at various dates 2018 – 2019 through <<https://www.oed.com/>>

Japanese:

'Jisho'. 2019. Ahlström, Kim, Ahlström, Miwa and Plummer, Andrew. Accessed at various dates 2019 through <<https://jisho.org/>>

Italian:

'Grande Dizionario Hoepli Italiano' 2018. Aldo Gabrielli. Hoepli. Accessed at various dates 2019 through <<https://dizionari.repubblica.it/italiano.html>>