



**HAL**  
open science

## English PhonologieS?

Quentin Dabouis, Pierre Fournier

► **To cite this version:**

Quentin Dabouis, Pierre Fournier. English PhonologieS?. Models and Modelisation in Linguistics, 2022. hal-03783885v1

**HAL Id: hal-03783885**

**<https://hal.science/hal-03783885v1>**

Submitted on 22 Sep 2022 (v1), last revised 9 Jan 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## English PhonologieS?

Quentin Dabouis<sup>a</sup> & Pierre Fournier<sup>b</sup>

- a. Université Clermont Auvergne, Laboratoire de Recherche sur le Langage (UR 999)
- b. Université Sorbonne Paris Nord, Pléiade (UR 7338)

### Résumé

Cette contribution propose que le système phonologique de l'anglais ne soit pas considéré comme un bloc uniforme mais au contraire comme un ensemble de sous-systèmes possédant des caractéristiques accentuelles, orthographiques, morphologiques, sémantiques ou encore segmentales spécifiques. Nous passons en revue une série de données montrant l'interconnexion de ces caractéristiques, ce qui indique que les sous-systèmes ainsi identifiés ne relèvent pas uniquement de la phonologie. Ces sous-systèmes sont associés à des parties du vocabulaire qui peuvent être analysées comme « étrangères » ou « savantes ». Nous montrons que le vocabulaire « étranger » a un comportement qui est en partie indépendant du système phonologique des langues d'emprunt. Bien qu'une formalisation du système proposé nous semble prématuré, des outils existants qui la permettraient lorsque d'autres études empiriques auront été menées pour affiner notre analyse sont discutés.

### Abstract

This study proposes that the English phonological system should not be considered as a uniform entity but as a set of subsystems with specific stress, orthographic, morphological, semantic or segmental features. We review a series of data showing the interconnection of these features, which shows that the subsystems we posit are not purely phonological. These subsystems are associated to parts of the vocabulary which can be perceived as “foreign” or “learned”. We show that “foreign” vocabulary has a phonological behaviour that is partly independent from that of donor languages. Although formalizing the system that we propose seems premature, we discuss existing tools which will allow it when further empirical studies have been conducted to refine our analysis.

### Introduction

Studies of English morphology or of spelling-to-sound relationships have often used different subsystems such as “native” (or “basic”), “non-native” (or “Latinate”), “French” or “Italian” to describe the different generalizations found in the language. In his foundational work on English spelling, which uses orthographic subsystems, Carney (1994, 98) asserts that “we should expect these orthographic subsystems to be connected with other aspects of English structure, particularly with phonology and word-formation”. However, such subsystems have rarely been invoked in analyses of English phonology. Hammond (1999, 284) claims that “in some languages, borrowed words exhibit rather different phonological patterns from those of native words, but this does not appear to be the case in the stress system of English”. Such an assumption, possibly shared by other researchers although they do not explicitly say so, has led loanwords to be frequently cited as examples of the stress rules of English. For instance, the words *Ticonderoga* and *Monongahela* are cited to illustrate the rules of secondary stress placement in many reference works on English stress such as Burzio (1994), Chomsky and Halle (1968), Hayes (1982) or Pater (2000).

However, others have suggested that English phonology, and especially stress placement, is a hybrid and dynamic system in which several logics are at work, and that these can be attributed to the fact that English historically had root-initial stress and has massively borrowed words from Romance languages such as Latin and French with stress further away from the left edge (J.-M. Fournier 2007; Minkova 2006). The necessity of distinguishing different parts of the vocabulary was also advocated for by Chomsky and Halle (1968, 373):

In English, for example, we have noted that it is necessary to classify many lexical items in terms of a feature that, roughly, distinguishes items of Germanic origin from other items; and for certain rules, such as Velar Softening, we need a further classification of the non-Germanic part of the vocabulary into items of Greek and Romance origin, roughly. This classification is functional in the

language and must be presumed to be represented in the internalized grammar. It is justified not by the historical development of the language but by the applicability of phonological and morphological rules.

We are not aware of many works on English phonology that have pursued these observations and none has sought to systematically describe the potential subsystems found in English. However, Pater (2005), states that “phonological processes and structures are often limited to a particular set of a language’s words, such as loanwords, Latinate words in English, or Yamato words in Japanese” and, in Pater (1994), he refers to different “classes of borrowings and names” among which he finds most words with light penult stress (e.g. *anténna*, *Alabáma*, *Botticélli*), a pattern that is otherwise highly exceptional. In diachronic studies of English, we can cite Minkova (1997), who argues that, “following the model proposed by Itô and Mester (1995[a]), the ME<sup>1</sup> lexicon can be stratified into a core consisting of native and already assimilated words for which all relevant constraints obtain, and increasingly peripheral lexical strata mapped like concentric circles around the core, for which the core constraints are progressively weakened. The boundaries between all strata allow fluctuation”.

Chomsky and Halle (1968) had also noted that this type of distinction is necessary in other languages such as Turkish or Modern Russian. Later studies have confirmed the necessity for languages such as Norwegian (Dresher 2013) or Japanese (Itô and Mester 1999) to separate the phonological processes accounting for the pronunciation of loanwords from the ones accounting for the rest of the lexicon. In Itô and Mester (1999), loanwords belong to foreign strata, which are distinct from native ones, as it appears that the pronunciation of loanwords cannot be accounted for by using the same mechanisms. In the paper in which they propose a taxonomy of loanword prosody, Davis, Tsujimura and Tu (2012) informally note that, in English, the preservation of the location of the source prominent syllable is not systematic and may depend on the source language: they note that many Japanese borrowings take penultimate stress regardless of the position of the accent in Japanese while many words borrowed from French retain word-final stress. This conclusion has later been confirmed by P. Fournier (2016b) with a large corpus based on the data of two reference pronouncing dictionaries.

The aim of this paper is to show that English phonology cannot be said to be constituted of a single monolithic system, but rather that it is made up of several subsystems which are, at least partly, related to etymology, and that these subsystems encompass more than is usually assumed in analyses of loanword phonology. To our knowledge, the main phonological phenomenon to function differently from one subsystem to another is stress, so this is what we will focus on first (section 1). However, it is possible that other phonological processes also display non-uniform behaviour when these subsystems are taken into consideration and so we might want to have criteria to identify the relevant parts of the vocabulary based on properties that are independent of phonology. Therefore, we will detail some of the properties of these subparts of the vocabulary and show that they have specific graphophonological, segmental, morphological and semantic properties (section 2). As some of the subsystems that we propose relate to different types of borrowed vocabulary, we will discuss the issue of the relationship between the pronunciations found in the source languages and those found in English. Then, we will turn to the evidence available regarding how speakers deal with these subsystems (section 3) and then we argue that the pronunciations found in English for loanwords cannot simply be analysed as imitations of the pronunciations found in the source languages but that the subsystems we suggest have gained a form of independence with regard to these languages (section 4). Finally, we will discuss the possible ways to formalize the overall system we posit here (section 5).

---

<sup>1</sup> ME stands for Middle English.

In this paper, we will posit the five following subsystems, to which we will refer using “§”, following Carney's notation (1994)<sup>2</sup>.

- §Core: most of the vocabulary, with regular spelling-to-sound correspondences, nonspecialised semantics and productive morphological processes (e.g. affixation by juxtaposition, compounding). It will sometimes be necessary to distinguish two subsystems within §Core, although the two are very much intertwined:
  - §Core-Native: words of Germanic origin, which constitute most of the basic vocabulary and are usually quite short.
  - §Core-Latinate: old borrowings from Latin or French and which may not be clearly identified as such. Many of them are quite frequent and words tend to be longer than those of §Core-Native.
- §French: words associated to French pronunciation (e.g. final stress) and spelling-to-sound correspondences. They often carry semantics associated to a form of “prestige”.
- §Foreign: words borrowed from or associated to Romance languages such as Italian, Spanish or Portuguese (to the exclusion of French) but it may also include words from other language families such as Russian or Japanese. These are massively stressed on their penultimate syllable and often refer to foreign cultures (e.g. food, clothes, traditions).
- §Learned: technical or scientific vocabulary, often borrowed directly or constructed from elements borrowed from Greek or Latin. These include neoclassical compounds.

The assumption according to which five subsystems could be attested will be developed and more details on the properties of these subsystems will be given in the following sections, with particular emphasis on systems other than §Core. Following Carney (1994), an additional §Exotic subsystem including borrowings from other languages (e.g. Irish, Scots, Arabic) could be added but will not be discussed here because it does not concern many words and their use is often marginal<sup>3</sup>. This paper focuses on British English, but we assume that the analysis can be extended to other varieties, with only minor differences.

## 1. Subsystems and stress

As noted in the introduction, previous literature is unclear regarding whether or not loanwords have distinct stress behaviours in English<sup>4</sup>. In this section, we argue that they do, based on evidence from previous studies, and we argue that we need more than a distinction between “native” and “Latinate” subsystems. One of the first works to sketch out such a system is Church (1985), in an attempt to model stress assignment for text-to-speech synthesis. After observing that loanwords have different stress behaviours in comparable weight conditions (which he assumes, following Chomsky and Halle (1968), to be a determining factor in stress assignment), he states: “I believe that speakers of English adopt what I like to call a pseudo-foreign accent. That is, when speakers want to communicate that a word is non-native, they modify certain parameters of the English stress rules in simple ways that produce bizarre “foreign sounding” outputs”. This leads him to design a statistical tri-gram model that assigns etymological distinctions based on the probability that three letter sequences are associated with a particular

---

<sup>2</sup> Following Aronoff (1976, 52), we assume that the diacritics which indicate whether a word is §Core or §French are properties of morphemes and not of words.

<sup>3</sup> For example, P. Fournier (2021a) studies Scottish Gaelic loanwords in which he finds that some spelling-to-sound correspondences are reproduced in English (e.g. <m> - /v/ or <io> - /i:/).

<sup>4</sup> Throughout this paper, we represent stress patterns using the following notation: /1/ or an acute accent represent primary stress, /2/ or a grave accent represent secondary stress and /0/ or no accent represent the absence of stress.

etymology, which can then be used for stress assignment. Let us now explore this idea in more detail.

The first observation regarding stress placement concerns suffix classes. It can be noted that most stress-affecting suffixes are Latinate (here, analysed as belonging to §Core-Latinate) while most stress-neutral suffixes are Germanic (§Core-Native). However, there are reasons to question the attribution of the stress properties of suffixes to different subsystems. First, all consonant-initial suffixes are usually stress-neutral (e.g. *-ful*, *-ness*, *-less*) while most vowel-initial suffixes are stress-affecting or mixed (e.g. *-ity*, *-ic*, *-able*). Therefore, the ability of a suffix to affect stress placement could be attributed to its segmental makeup (as is proposed in analyses which oppose cohesive and non-cohesive suffixes; see e.g. Raffelsiefen (2015)), and this segmental makeup just happens to be related to etymology: Germanic suffixes usually begin with a consonant while Latinate suffixes usually begin with a vowel. Second, the distribution of stress properties for suffixes does not always fit the description given above. Indeed, there are Latinate suffixes which are generally quite neutral, although they often exhibit a mixed behaviour (e.g. *-ize*, *-able*, *-ism*). The occasional cases in which the suffix appears to be stress-affecting can generally be accounted for because the relationship between the base and the derivative is no longer semantically transparent (e.g. *canál* → *cánalize*; *compáre* → *cómparable*) and such cases can occur in suffixed neoclassical compounds (e.g. *épilogue* → *épilogize*; *díalogue* → *diáloguist*), even though such stress shifts do not necessarily correlate with semantic drift in the latter cases<sup>5</sup>. There is also one case of a Germanic consonant-initial suffix which is evolving towards stress-affecting behaviour: *-ly* now often shifts stress to the antepenultimate syllable in *-arily* words (e.g. *nécessary* → *nècessáryly*), as discussed by Trevian (2007)<sup>6</sup>. Therefore, it is not certain that the stress properties of suffixes (stress-affecting vs. stress-neutral) should be attributed to the difference between §Core-Native and §Core-Latinate. Historically, there was indeed a difference as Germanic languages like Old English had root-initial stress while Latinate languages usually have stress closer to the right edge of the word. However, it is possible that this difference is not relevant anymore in contemporary English considering that, as we will see for other properties, the two subsystems in §Core have fused to a large extent.

Let us now consider final stress in nouns. In English<sup>7</sup>, final stress is quite uncommon in nouns and is considered the exception rather than the rule. Generative approaches usually assume that the last syllable of nouns is extrametrical, i.e. it is invisible to stress-assigning processes and therefore may not receive primary stress (Hayes 1982). Many of the nouns that do have final stress are in fact borrowings, often from French (e.g. *champágne*, *cuisíne*, *fatígue*, *màquillage*, *pèrsonnèl*, *petíte*, *routíne*) but occasionally from other languages (e.g. *amír*, *canóe*, *cartél*, *imám*, *pàrmesán*). Some of these words are even organised into micro-paradigms which resist stress retraction, often sharing a suffix (e.g. *àmatéur*, *danséur*, *rèstauratéur*, *arcáde*, *bàrricáde*, *grenáde*, *paráde*, *millionáire*, *quèstionnáire*, *grotésque*, *hùmorésque*, *pìcturésque*). Therefore, we could claim that final stress is regular in §French nouns<sup>8</sup> while it is dispreferred for other nouns.

---

<sup>5</sup> For discussions on mixed affixes, see Bermúdez-Otero (2018), Trevian (2007) or Giegerich (1999).

<sup>6</sup> See also Castanier (2016) for a study of the historical development of this stress behaviour.

<sup>7</sup> British English has long been considered as the reference variety when dealing with the pronunciation of English. However, phonological studies examining other varieties have been more and more present in the literature for the last two decades. Nevertheless, it appears that stress variation across varieties of English is restricted to few isolated words (Martin 2011), except for French loanwords. Indeed, the stressing of French loanwords is different in British English and in American English because such parameters as noun category (i.e. proper name vs. noun) or morphology do not have the same impact in both varieties (P. Fournier 2016a).

<sup>8</sup> The pronunciation of proper nouns is very complex and often ignored in reference studies. However, this distinction could turn out to be significant when applied to the stressing of French loanwords. P. Fournier (2016a) shows that the proper noun / common noun distinction is not significant in American English because words

In a study of the role of syllable weight in stress assignment in English, J.-M. Fournier (2010a) found that the impact of a prefinal consonant cluster on stress<sup>9</sup> differs depending on the type of vocabulary concerned. If one puts aside words with adjectival suffixes, whose stress patterns may be attributed to effects of the suffixes (e.g. *ascendant*, *persistant*, *defensible*, *attractive*, *incurvate*, *enormous*), two main classes of words can be identified. The first one is made up of words borrowed from Romance languages such as Italian, Spanish or Portuguese (e.g. *anacóna*, *extravaganza*, *dilettante*, *concerto*) or from Modern Latin (e.g. *propaganda*, *enigma*, *memo*, *alumnus*), and both have almost systematically penultimate stress. Words with orthographic consonant geminates can be associated to this first class (e.g. *salmonella*, *dilemma*, *antenna*, *risotto*, *tobacco*, *espresso*). However, the second group is made up of words which cannot be clearly identified as “foreign” and displays a lot more stress heterogeneity than the first group: 93 words have antepenultimate stress (e.g. *discipline*, *sacristan*, *calendar*, *lavender*, *minister*, *industry*, *interval*) while only 53 have penultimate stress (e.g. *advantage*, *intestine*, *adventure*, *epaulment*, *coriander*, *semester*, *elixir*). J.-M. Fournier therefore claims that a prefinal consonant cluster is only decisive for stress assignment in borrowed or learned vocabulary. Moreover, Guierre (1979, 544–50) introduced a rule for “Italian” words. This rule states that words of the form <-VCV#> whose prefinal consonant is a dental obstruent and whose final vowel is a pronounced vowel other than <y> have penultimate stress (e.g. *casino*, *libido*, *medusa*, *potato*). Words with another prefinal consonant in that configuration are approximately equally divided between penultimate and antepenultimate stress (e.g. *gorgonzola*, *safari*, *impetigo* vs. *formula*, *camera*, *indigo*). This raises a question: does the specific segmental configuration of Guierre’s “Italian” words define these words as “foreign” or are there other factors that define these words as “foreign”, assuming that only foreign words follow this rule, which is conditioned by segmental structure? It is quite possible that both possibilities co-exist and interact. With the subsystems that we assume here, we can say that sensitivity to a prefinal consonant cluster is shared by §Foreign and §Learned, but that they differ with regard to words with a single prefinal consonant. §Foreign words tend to be stressed on the penultimate, especially in the environment of Guierre’s “Italian” words rule. As for §Learned words, we do not know of any study that has sought to specifically analyse their stress pattern, but two observations can be made. First, there are a few technical suffixes associated to penultimate stress (e.g. *-osis*, *-itis*, *-oma*; J.-M. Fournier 2010a, 29) and, in words free of morphological influence, J.-M. Fournier (2010a, 70–71) notes that many words borrowed from Latin have penultimate stress (e.g. *asylum*, *decorum*, *erratum*, *hiatus*, *papyrus*, *verbaticum*), but his list of words with antepenultimate stress also contains a few §Learned words (e.g. *asparagus*, *basilica*, *lexicon*, *maximum*, *specimen*, *syllabus*), so we are not in a position to say what the “regular” stress pattern is for §Learned words with a single prefinal consonant<sup>10</sup>.

Among §Learned words, there is a class of words that have a clearly distinct stress behaviour: neoclassical compounds, i.e. words made up of Greek or occasionally Latin roots such as *bio-*, *-graph* or *-scope*. This class is a well-known issue in morphology, but the stress properties of neoclassical compounds have been described as their “most striking” property (Guierre 1979, 740) and the “only [criterion] which is really clear<sup>11</sup>” to define this class of

---

coming from French are massively stressed on the final syllable, regardless of the category they may belong to. However, in British English, proper nouns coming from French exhibit more non-final stress patterns than common nouns.

<sup>9</sup> In his approach, a functional consonant cluster roughly corresponds to a cluster that would be syllabified heterosyllabically. Therefore, words with a prefinal consonant cluster are words that are usually analysed as having a closed penultimate syllable.

<sup>10</sup> Some will assume that the reason why words such as *asylum* have penultimate stress is that they have a heavy penultimate syllable. This is not the position adopted here: we assume that stress may determine the distribution of vowels, but that vowels may not determine the position of stress.

<sup>11</sup> In French: “seul [critère] qui soit vraiment clair” (our translation).

words and distinguish it from affixes (Tournier 1985, 92). Their stress properties can be described as follows:

- In non-suffixed compounds, primary stress is assigned to the first element as if it were an independent word. As Trevian (2015, 430–34) shows, the first element usually has penultimate stress if it has a prefinal consonant cluster (e.g. *larýngoscope*, *eléctrotype*) or if the onset of the last syllable contains a consonant other than /k, g, l, r/ (e.g. *adreno-*, *carcino-*, *dynamo-*, *germano-*). Otherwise, it receives antepenultimate stress. This may lead to pre-antepenultimate stress if the first element is trisyllabic and has initial stress (e.g. *áudiophile*, *bíbliophile*, *cárdiograph*, *héterodox*, *ídeophone*, *méritocrat*, *rádiogram*, *sóciopath*, *stéreotype*; J.-M. Fournier 2010a, 75–76)<sup>12</sup>. Furthermore, Guierre (1979, 741) shows that in non-suffixed compounds with monosyllabic final roots, primary stress is generally located on the antepenultimate syllable, whatever the segmental structure of the penult, i.e. regardless of whether or not it is followed by a consonant cluster (e.g. *báthyscaph*, *stéthoscope*, *ápophthegm*, *gýrostat*, *thérmostat*, *áutopsy*). In that regard, these words are distinct from §Learned words which are not neoclassical compounds and from §Foreign words.
- In suffixed words, two generalizations can be made:
  - many suffixes that are usually neutral or mixed become stress-affecting, including learned suffixes such as *-is*, *-on* or *-us*, imposing penultimate stress on words with a prefinal consonant cluster and antepenultimate stress elsewhere (e.g. *búreaucrat* → *buréaucracy*; *télegraph* → *telégraphist*; *phótoph* → *photógrapher*; J.-M. Fournier 2010a, 76–77)<sup>13</sup>.
  - certain roots, when found in the second position of a suffixed neoclassical compound, attract stress on themselves, while others let stress fall on the antepenultimate syllable (e.g. *èpidúral* ~ *epíscopal*; *hèxahédral* ~ *hexágonal*; *ántidótal* ~ *antíphonal*; Dabouis 2016b; P. Fournier 2013, 2017).
- Guierre (1979, 740) claims that neoclassical roots are accentually invariant. As pointed out by Dabouis (2016a, 246), this could potentially override weak stress preservation. He only provides anecdotal evidence: *làryngólogy* → *larýngológical*. In this case, we would expect the initial secondary stress in *laryngology* to be preserved in *laryngological*, but instead we get a stress shift to the second syllable. This shift could be attributed to the fact that the default stress for the root *laryngo-* is on its second syllable (e.g. *larýngoscope*, *larýngograph* cp. *céphalograph*, *métallograph*, *gálvanoscope*, *láparoscope*). This is further supported by the fact that Dabouis (2016a) finds no uncontroversial cases of stress preservation failure on the first syllable. However, more empirical work is still needed to test this claim.

The generalizations we have detailed in this section for the different subsystems are distinct from those which regulate stress in §Core. In §Core, simplex words generally follow the Normal Stress Rule: disyllabic words get initial stress and longer words get antepenultimate stress. As §Core-Native words are generally quite short, this generally corresponds to word-initial stress. Semantically opaque prefixed words get primary stress on their root in words

---

<sup>12</sup> A notable exception is the case of Neo-Latin compounds, which can be recognized by their linking *-i-*. These usually have antepenultimate stress, even if that leads to mismatches with transparent bases (e.g. *humídístat*, *inséctífuge*, *orgánígram*; Trevian 2015, 431).

<sup>13</sup> Two limitations of this generalization can be mentioned. First, Trevian (2007) nuances that claim for *-ist* and *-ism* and brings forward several counterexamples (e.g. *párallel* → *párallelism*; *cátalogue* → *cátaloguist*). Second, J.-M. Fournier (2010a, 77) notes that certain learned roots “neutralize” the suffixes that are usually neutral or mixed (e.g. *óligarch(y)*, *héterodox(y)*, *eléctrotype(ist)*).

which are not nouns, whether they are §Core-Native (e.g. *believe*, *forget*, *tomorrow*, *understand*, *withhold*) or §Core-Latinate (e.g. *contain*, *deflect*, *extinct*, *rejoice*, *submit*). This is what J.-M. Fournier (2007) calls the Germanic Law<sup>14</sup>. Suffixes are quite generally stress-neutral, except the handful of stress-affecting suffixes, mainly found in §Core-Latinate (e.g. *-ity*, *-ic*, *-ion*). Finally, compound words quite generally belong to §Core, and are regulated by the Normal Compound Stress principle, which assigns primary stress to the first element of the compound (e.g. *áircraft*, *bédroom*, *cóuntrieside*, *éverything*, *mótorway*, *páperwork*, *wárdrobe*).

Let us finish with one more remark on secondary stress. The specificities of foreign vocabulary appear to concern only primary stress (to the possible exception of neoclassical compounds). Indeed, in a large study on secondary stress in British English, Dabouis (2016b) examined foreign words and learned words (words containing neoclassical roots were excluded) separately to control for potential differences between the different types of vocabulary and found no significant differences between them with regard to secondary stress placement.

## 2. Other properties of the subsystems

### 2.1. Orthography and phonology

Several studies of English spelling-to-sound correspondences have expressed the necessity to distinguish subsystems. The first we are aware of is Albrow (1972), who distinguishes three systems, which are related to, but not defined exclusively by, etymology. These were later expanded and refined by Carney (1994) who, as discussed in the introduction, uses subsystems such as §Basic, §French or §Latinate. In the Guierian School, a research tradition mainly practised in France (Dabouis et al., forthcoming) which has also studied spelling-to-sound correspondences in detail, a few “foreign” vowels have been identified, although subsystems have not been posited as such. In this section, we will discuss vowels first and then consonants. Most of the examples are taken from Carney (1994), Deschamps (1994) or J.-M. Fournier (2010b), occasionally complemented by manual searches in Wells ([1992] 2008).

Guierre (1979), Deschamps (1994) and J.-M. Fournier (2010b) have developed a system of reading rules, mainly focused on stressed vowels. Therefore, the perspective is that of a reading grammar. In their system, vowels can be read in different ways, called “values” depending on the orthographic and/or phonetic context in which they are found (see Dabouis (forthcoming) for an overview). These rules often parallel phonological rules put forward by Chomsky and Halle (1968), although Guierre (1979) argues that the grammar of the reader is the most efficient, i.e. that including spelling makes for a more efficient system of rules.

This system reveals a number of “anomalies” found in foreign parts of the vocabulary, with pronunciations which do not fit with the general system, but rather seek to reproduce the phoneme found in the source language with the closest English phoneme, through a principle of articulatory proximity (Tournier 1993, 149). Such exceptional vowels have been called “foreign free vowels”. The three most common ones, along with the expected regular free vowels are shown in Figure 1.

---

<sup>14</sup> This is in line with analyses assuming that semantically opaque morphological structure may impact phonology (see Dabouis (2017) for arguments).



Figure 1: The main three foreign free vowels<sup>15</sup>

Free vowel	Spelling	Foreign free vowel	Examples
/eɪ/	<a>	/ɑ:/	<i>armada, banana, tomato</i>
/i:/	<e>	/eɪ/	<i>cliché, peso, sake</i>
/aɪ/	<i>	/i:/	<i>ballerina, bikini, mosquito</i>

From J.-M. Fournier (2010, 113)

These vowels are often said to be unpredictable, although they tend to occur in contexts where free vowels are expected, which J.-M. Fournier (2010b, 114) calls “host structure[s]” (Fr. “structure d’accueil”). Such vowels do sometimes occur in environments where other values are expected (e.g. *bouquet, motif*). Some variation does occur, sometimes across varieties, as in the examples in (1).

- (1) *albino* /æɪ'bi:nəʊ/ (GB) ~ /æɪ'bainəʊ/ (US)  
*tomato* /tə'mɑ:təʊ/ (GB) ~ /tə'meitəʊ/ (US)

The correspondences shown in Figure 1 are commonly found in §French words (e.g. *cliché, critique, montage*), in §Foreign words (e.g. *bikini, koala, sombrero*) and occasionally in §Learned words for <a> (e.g. *apparatus, diorama, erratum*). Note that there are no foreign free vowels for <o> and <u>, quite likely because their regular free vowels, /əʊ/ and /u:/ are already articulatorily close to the pronunciation of these orthographic vowels in the source languages. However, some cases of free vowels found in contexts in which other values are expected in foreign words with <o><sup>16</sup> suggest that even that orthographic vowel may behave differently in foreign words. Examples are shown in (2).

- (2) a. *coterie, omega, sobriquet*  
b. *omerta, origami, rodомontade*

The words in (2a) are given by J.-M. Fournier (2010b, 134) as exceptions to Luick’s rule, which predicts checked vowels in antepenultimate stressed vowels followed by a single consonant. Note that these three words may all have a foreign free vowel of the types given in Figure 1 elsewhere in the word<sup>17</sup>. The words in (2b) are examples of words in which <o> has secondary stress and which were found by Dabouis (2016a, 2018) also to resist Luick’s rule (along with expected values for *origami* and *rodомontade*). Like those in (2a), these may all have a foreign free vowel elsewhere in the word. Dabouis suggests that these exceptions are due to the foreign character of these words.

However, the diversity of foreign vowels and the contexts they appear in are more diverse than what is shown in Figure 1. Consider the cases in (3), where the correspondences listed in (3a) are typically found in §French words while those in (3b) are found in §Learned words.

<sup>15</sup> These can also be coloured by <r>, as in *bolero, lira, sombrero, Valkyrie*, although this does not affect /ɑ:/, e.g. *curare, safari, tiara* (J.-M. Fournier 2010b, 130).

<sup>16</sup> We do not know of any comparable cases with <u>. This is probably due to the fact that <u>, contrary to other vowels, is always free in the <(-)uCV(-)> context when it is stressed, and that this makes it almost impossible to find cases with an unexpected free vowel.

<sup>17</sup> J.-M. Fournier also gives words which display the same type of exception for words which cannot be claimed to be perceived as foreign (e.g. *codify, glorify, motivate, notify, potentate, probity, procreate*). These can often be attributed to exceptional vowel preservation from the base (cp. *code, glory, motive, note, potent*).

- (3) a. <ou> - /u: - uə/: *bijou, bourgeois, courgette, rendez-vous*  
 <eu> - /ɜ:/: *danseur/danseuse, effleurage, saboteur*  
 <(e)au> - /əʊ/: *auberge, bureau, château, chauffeur*  
 <oi, oy> - /wa:/: *boudoir, bourgeois, memoir, reservoir*  
 <é(e), ê, è> - /eɪ - eə/: *négligée, déjà vu, fiancé, résumé*  
 <i> - /æ/: *distingué, meringue, timbre*  
 <a, e> - /ɒ/: *blanquette, entrecote, langoustine*
- b. <ae, oe> - /i:/: *aesthetic, amoeba, Caesar, foetus, oecumenical*  
 <eu> - /ju:/: *eulogy, pneumonia, therapeutic*

Let us now turn to consonants. The information here is taken from Deschamps (1994), Carney (1994) and Trevian (1993, 2003). The most common “foreign” correspondences for consonants are listed in (4).

- (4) a. §French  
 <ch> - /ʃ/: *douche, ruche, niche, gauche*  
 <g> - /ʒ/: *camouflage, montage, prestige, refuge* (often variation with /dʒ/)  
 <qu> - /k/: *bouquet, critique, etiquette, mannequin, piquant*<sup>18</sup>  
 <gn> - /nj/: *cognac, mignonette, poignant, vignette*
- b. §Foreign  
 <c(c)> - /tʃ/: *ciabatta, cello, concerto, Boccaccio, Puccini, Riccio, capuccino*  
 <sch> - /ʃ/: *schadenfreude, schnaps, schmuck, Schrodinger*  
 <z> - /ts/: *nazi, panzer, zollverein*
- c. §Learned  
 <ph> - /f/: *apocrypha, diaphragm, epitaph, naphtha, phantom, sphere*  
 <ch> - /k/: *archaism, chaos, echo, matriarch, orchestra, stomach, technical*  
 <rh> - /r/: *rhapsody, rhetoric, rhizome, rhododendron, rhythm*

To these correspondences, we can add the cases for which a final orthographic consonant is silent. Guierre (1979, 788) lists 64 such words (out of 14,700 words ending in <-C>) and all of them can be analysed as §French (e.g. *coup, dossier, bourgeois, precis, éclat, buffet, argot, debut*).

Finally, the main distinction of which we are aware between items in §Core-Native and §Core-Latinate is the behaviour of <c> and <g>. As pointed out by Chomsky and Halle (1968, 373), the rule of so-called “Velar Softening” is restricted to Latinate words. This rule, which they assumed to be a phonological one, was rejected as such (Hyman 1975; Kaye 1995) and was later analysed as a spelling-to-sound rule (Montgomery 2001; Raffelsiefen 1993, 73–74), which can be formulated as in (5).

- (5) <c> → /s/ } -<i>  
 <g> → /dʒ/ } -<e>  
 } -<y>

Although he does not explicitly link it to his §Basic and §Latinate subsystems, Carney (1994) notes that exceptions to the correspondences in (5) generally have high frequencies, which is a clue that they are probably §Core-Native. This is confirmed by the examples he gives (e.g. *gear, get, giggle, gimmick, give*) which are quite generally §Core-Native.

<sup>18</sup> This is to be opposed to the expected /kw/ (e.g. *esquire, conquest, liquid, quack, square*) but it can also be found in non §French words (e.g. *conquer, exchequer, liquor, queue*).

## 2.2. Segmental structure

The first representative criterion of segmental structure is word length, which tends to differ depending on the origin of words. Plag (1999, 59–60) asserts that native (§Core-Native) and Latinate words (§Core-Latinate) often differ in their phonological structure. Thus, it has been frequently pointed out that Latinate stems tend to be polysyllabic with stress patterns different from those of Germanic stems which tend to be either monosyllabic, or, if polysyllabic, to bear primary stress on the first syllable. §Foreign words also tend to be polysyllabic whereas there are many monosyllabic items among §Core-Native words, which is due to the structural properties of the Germanic lexicon<sup>19</sup>.

Moreover, §Foreign words often exhibit <-VCV#> structures, which are inherited from the syllabic structure of Romance languages. This orthographic structure is often attested in Japanese loanwords as well. This structure is so much correlated with the penultimate stress pattern that it has been considered to be a determining parameter in stress assignment along with segmental restrictions (see the rule on “Italian” words discussed in section 1). Furthermore, as noted by J.-M. Fournier (2010b, 114), in §Foreign words ending in <-VCV#>, the final segment is often a monographic vowel (other than <y>) and which is necessarily pronounced (this excludes final mute <e#>).

§Learned words include neoclassical compounds. These words exhibit two specific segmental features. The first one is the orthographic structure of what has been labelled “B roots” by Trevian (2003). The morphological structure of neoclassical compounds can be analysed as AB(C), where A stands for the first root, B for the second one and C for an optional suffix. According to Trevian, (2003, 183), B roots often have <-CVC-> structures (examples: -*vor-*, -*log-*, -*phil-*). This segmental alternation between consonants and vowels in the prefinal syllables of words interestingly reminds one of the segmental properties of §Foreign words. Actually, this regular C/V alternation might be a good indicator for native speakers to identify the supposedly foreign origin of §Foreign and §Learned items. The second segmental characteristic of neoclassical compounds is the presence of a medial vowel between the two roots: <i> for compounds of Latinate origin and <o> for compounds of Greek origin.

Finally, §French words show specific phonotactics for the distribution of /ʒ/. Indeed, /ʒ/ is only attested in initial and final positions in French loanwords (e.g. *genre*, *jeté*, *julienne*, *bon voyage* (and a few words ending in *-age*)).

## 2.3. Morphology

In terms of morphology, we will discuss four types of characteristics: specific formatives, specific morphological processes, selectional restrictions and productivity. We will focus on characteristics of the subsystems other than §Core, which makes up the main part of the lexicon, and to which most formatives and productive word-formation processes (e.g. affixation by juxtaposition, compounding, conversion) are associated.

Let us first consider formatives. There are certain formatives that have been borrowed from other languages and can therefore be used as cues to identify a word as being a part of a subsystem. This is the case for a few French suffixes such as *-ette*, *-esque*, *-aire*, although some of them seem to be undergoing a form of integration into §Core. This is the case for *-ade*, which is almost systematically associated with final stress but whose pronunciation oscillates between /-eid/ and /-ɑ:d/ (e.g. *charade*, *façade*, *esplanade*). Another interesting case is *-age*, as words that contain it have often undergone stress retraction but, at the same time, a “denativization” of the vowel of the suffix of the form /-idʒ/ → /-ɑ:(d)ʒ/, as in *espionage* (Castanier 2016, 225).

---

<sup>19</sup> As pointed out by Minkova (2006), foreign vocabulary became predominant in polysyllabic words between 1500 and 1700 A.D. because of the massive influx of words from Latin and French and because many native words became shorter due to the attrition and loss of final unstressed vowels.

Another series of suffixes have been borrowed from Latin and can be associated to §Learned, such as *-(i)a*, *-i*, *-is*, *-on*, *-um* or *-us*.

A number of §Learned words contain neoclassical roots (e.g. *bio*, *laryngo*, *type*, *gen*, *arch*; also called “quasi-morphemes” or “combining forms”), which are mostly borrowed from Greek. These can sometimes be difficult to distinguish from affixes or from bound roots (Tournier 1985, 88 ff.; Bauer, Lieber and Plag 2013, 18), but they are often analysed as constituting a separate class. As pointed out by Bauer, Lieber and Plag (2013, 455), neoclassical roots may attach to bound roots (e.g. *glaciology*, *scientology*), words (e.g. *hydro-electric*, *morpho-syntax*) or another neoclassical root (e.g. *biology*, *telegraph*). We can also add that they can sometimes be found with only an affix (e.g. *electron* can be argued to be *electro-* + *-on*). One of the properties that distinguish them from affixes is that they may appear before or after another root (e.g. ***graphology***, ***telegraph***; ***psychology***, ***logomachy***), for at least a quarter of them according to Trevian (2003, 176), which leads Tournier (1985, 89) to propose that they “have the *potential* capacity to be placed on the left or on the right, and that there is no reason for this possibility to be realised – or even realisable for all [neoclassical roots]”<sup>20</sup>. Nonetheless, some elements tend to occur on the left while others tend to occur on the right, but Tournier argues that this is rather a question of probability than of possibility. Therefore, neoclassical roots are a special type of formative and they are used in specific word-formation processes.

Guierre (1988) notes that base truncation, i.e. the loss of the rhyme of the last syllable of a base and its replacement by a suffix, is almost systematic for words that end in a pronounced vowel monograph that is not <y>, which, as we have seen, are typical structures of §Foreign or §Learned words (e.g. *Genev/a*, *-an*; *Chin/a*, *-ese*; *alkal/i*, *-oid*; *infern/o*, *-al*). However, he finds that this is not true of suffixes beginning with <i>, which sometimes trigger base truncation and sometimes do not (e.g. *Ved/a*, *-ic* but *formul/a*, *-aic*; *esperant/o*, *-ist* but *jing/o*, *-oism*). He also notes that base truncation is systematic, even with suffixes beginning with <i>, in words ending in *-us* or *-um*, which we have identified as typical §Learned suffixes (e.g. *foc/us*, *-al*; *fung/us*, *-oid*; *vel/um*, *-ar*; *narciss/us*, *-ic*). These behaviours are quite distinct from what can be found in §Core derived words, in which suffixes almost never trigger base truncation<sup>21</sup>.

In the literature on English morphology, we often find a distinction between “native” and “non-native” (or “Latinate”) formatives, especially in studies on the selectional restriction of affixes. Recent research confirms that this distinction is indeed relevant, although it is not as strong as some have previously assumed:

We note that while there are non-native suffixes that only attach to simplex bases, there are no native suffixes that are restricted with regard to the etymological origin of the simplex bases they attach to, although there are a few that attach only to simplex bases.

All of the native prefixes attach indiscriminately to both native and non-native bases, as do the vast majority of non-native prefixes. There are only five non-native prefixes that never seem to attach to native bases (*a-*, *circum-*, *cis-*, *NEGin-*, *mal-*) and eight that attach to native bases only rarely (*contra-*, *hemi-*, *peri-*, *pico-*, *supra-*, *tera-*, *trans-*, *uni-*). It is notable that none of these prefixes is particularly productive in contemporary English (Bauer, Lieber and Plag 2013, 611).

This supports the idea that a distinction between §Core-Native and §Core-Latinate is still required, although it shows that the two systems are not completely separate.

Finally, it is often assumed that native affixes are more productive than non-native affixes, especially because native suffixes are usually consonant-initial while non-native

<sup>20</sup> In French: “ont la capacité *potentielle* d’être placés à gauche ou à droite, et qu’il n’y a pas de raison pour que cette éventualité soit réalisée – voire réalisable pour *tous* les q.m.” (our translation).

<sup>21</sup> A notable exception, discussed by Guierre and many others, is that of base truncation in *-ate* verbs to which *-able* is added (e.g. *demonstr/ate*, *-able*; *navig/ate*, *-able*).

suffixes are generally vowel-initial, which tends to make the former more “parsable” than the latter (Hay 2003). However, it has been shown that phonotactics alone cannot predict productivity (Plag 2002) and that other factors such as usefulness and parsability need to be taken into account (Hay and Baayen 2003). In certain cases of competing affixes, the productivity of affixes cannot be accounted for using the native/non-native distinction. For example, in her detailed study of *-ity* and *-ness*, Arndt-Lappe (2014) observes that “differences are more adequately captured in terms of the scope of domains in which the two suffixes are productive” and shows that *-ity* has become more productive in certain domains over the past three centuries.

#### 2.4. Semantics

One of the defining criteria of what might be perceived as foreign is of course semantics. As early as 1933, Bloomfield uses the expression “culture loans” (1933, 458) to refer to the semantic fields covered by loanwords. We can expect words which refer to cultural habits, clothes or food that are not typical of English-speaking countries to be perceived as foreign. A few examples are shown in (6), where the definitions are from Dictionary.com.

- (6) *bolero* “a lively Spanish dance in triple meter.”  
*emir* “a chieftain, prince, commander, or head of state in some Islamic countries.”  
*kimono* “a loose, wide-sleeved robe, fastened at the waist with a wide sash, characteristic of Japanese costume.”  
*maquis* “the French underground movement, or Resistance, that combatted the Nazis in World War II.”  
*spaghetti* “a white, starchy pasta of Italian origin that is made in the form of long strings, boiled, and served with any of a variety of meat, tomato, or other sauces.”  
*tokay* “an aromatic wine made from Furmint grapes grown in the district surrounding Tokay, a town in NE Hungary.”

Moreover, foreign words are sometimes very present in certain lexical fields. For example, Ryan (2017, 308) notes that “coffee has a lot of §Italian words: *espresso*, *macchiato*, *cappuccino*, *latte*”. The same is true of musical terminology (e.g. *piano*, *forte*, *crescendo*, *allegro*). We can find similar things for §French words, for example for food and cooking (e.g. *cuisine*, *casserole*, *mayonnaise*, *puree*, *soufflé*). Tadmor (2009, 54) identifies the following semantic fields as the most affected ones by the borrowing process: religion and belief, clothing and grooming, the house, law, social and political relations, agriculture and vegetation and, finally, food and drink. Likewise, §Learned words, are generally technical or scientific vocabulary, although a few of them have come into common usage (e.g. *democracy*, *economy*, *photograph*, *suicide*). Coates (1999, 27) describes neoclassical roots as “a kind of parallel vocabulary, a resource for the creation of words for special purposes, especially scientific and technical”.

Let us underline that semantics seems to be particularly important for French borrowings. Indeed, these borrowings often have a stylistic function in that they express “a certain image of France, its culture, of its ways of life; a certain ‘French way of life’<sup>22</sup>” (Chiról 1976, 149). Chadelat (2000) speaks of the notions of “prestige” to describe the connotations often associated to French borrowings (Haspelmath (2009, 35) uses the expression “attitudinal factors” to refer to these extralinguistic parameters). He also notes that the loss of final stress in such words through stress retraction goes along with the loss of these connotations and leaves simply a referential function (Chadelat 2000, 194). Castanier’s (2016) work on the evolution of

---

<sup>22</sup> In French “une certaine image de la France, de sa culture, de ses modes de vie; un certain « [F]rench way of life »” (our translation).

stress placement in the last three centuries seems to overall confirm this observation. For example, in French borrowings ending with *-age*, the most frequent and most “useful” words (in certain idiolects at least) tend to undergo stress retraction (e.g. *fuselage*, *sabotage*, *entourage*) while words that are less frequent, less useful and which still have connotations of “prestige”, such as *maquillage*, preserve final stress.

All of this is to be distinguished from §Core vocabulary. Everyday vocabulary is mainly §Core-Native, as can be seen if we look at the most frequent word-forms in SUBTLEX-UK (Van Heuven et al. 2014) for example, even excluding function words (e.g. *all*, *just*, *think*, *get*, *going*, *got*, *out*, *here*, *up*, *now*). However, many §Core-Latinate words are also frequently used (e.g. *question*, *moment*, *important*, *second*, *problem*).

## 2.5. Summary

The list of properties which we have presented so far, which is summed up in Figure 2, should not be seen as a final account of the properties of the subsystems we propose. As our claim has not been the main focus of any large empirical study, what we have presented so far should be refined and completed by future research. One possible additional phenomenon to include is vowel reduction, as Guierre (1979, 746) shows that, in neoclassical compounds, the final roots are generally not reduced when they are unstressed (e.g. *gazogene*, *xylophone*, *paralyse*, *metronome*, *gramophone*, *anglophile*, *periscope*, *antidote*, *oligarch*, *semaphore*), except for a few cases, which turn out to be frequently used in English (e.g. *monarch*, *oxygen*, *symptom*, *pentagon*). Moreover, Fidelholtz (1975) found that vowel reduction applied more frequently to words with a higher frequency. It is possible that part of this phenomenon can be explained by the fact that words from §Foreign, §French or §Learned have generally lower frequencies than the words of §Core and, for some reason, resist vowel reduction more than the latter. Exploratory work which we are currently conducting suggests that there are indeed differences between §Core words and words from the other subsystems regarding vowel reduction, which seems to confirm the idea: the former reduce more than the latter, even when word frequency (along with other parameters) is controlled for.

Figure 2: Summary of the properties of the different subsystems

		Stress			Graphophonology		Morphology	Semantics	Segmental structure
\$Core	§Native	Normal Stress Rule	Neutral suffixes?	Regular secondary stress rules	Basic correspondences	No Velar Softening	Affixation by juxtaposition Compounding Conversion	Basic vocabulary	Short words
	§Latinize	Germanic Law	Stress-affecting or mixed suffixes?			Velar Softening	Some selectional restrictions for affixes Base truncation in <i>-able</i> derivatives ← <i>-ate</i> bases	Common vocabulary	Polysyllabic words
\$French		/-1/ Some stress-affecting suffixes (e.g. <i>-ette, -aire, -eur</i> )			<a> - /a:/ <e> - /e:/ <i> - /i:/	Specific correspondences (3a) & (4a)	Specific suffixes (e.g. <i>-ette, -aire</i> )	« Prestige »	Distribution of /ʒ/
\$Foreign		/-10/ (especially “Italian” words)				Silent final consonants			
\$Learned		/-10/ if prefinal consonant cluster Otherwise /10/ or /-100/? Specific stress rules for neoclassical compounds	Specificities for secondary stress placement in suffixed neoclassical compounds?	<a> - /a:/ (marginal)	<ae, oe> - /i:/ <eu> - /ju:/ <ph> - /f/ <ch> - /k/	Specific suffixes (e.g. <i>-itis, -osis -us, -on</i> ) Base truncation more common Neoclassical compounding	Technical, scientific or learned vocabulary	Polysyllabic words CVC Structure for B roots in neoclassical compounds	

### 3. Speakers and the lexicon

Studies in English phonology rely either on oral or dictionary-based data. However, the observation of statistical regularities based on dictionary-based corpora has to be distinguished from what native speakers can actually produce. The two dimensions are closely linked but fall within two distinct notions: language and speech. Problems can be underlined with both approaches. Oral corpora are generally based on the productions of very few speakers and cannot be said to be truly representative of a given variety of English. The same is true for dictionary-based analyses which can be argued to describe an “artificial idiolect” of English (Collie 2007, 116–18) with little or no reference to actual speakers. We believe that both approaches could be complementary when dealing with loanwords. Loanwords are borrowed from source languages thanks to speakers, who may or may not be bilingual. The first step is oral, and, only later, once used and understood by a large community of speakers, are items integrated into dictionaries and considered as institutionalised. However, some criticisms might focus on “naive” speakers and their possible inability to pronounce loanwords in the way that is advocated in dictionaries, especially loanwords which could be classified as “exotic”. This cannot be attributed to the lack of “realism” of dictionaries but rather to the tendency that native speakers have, when facing unknown items, to use the generalizations accounting for the pronunciation of the words in their own lexicon. And the pronunciation of loanwords is a perfect example of conflicts between “learned” speakers (not to use “bilingual” which would probably be an overstatement, because speakers with notions in foreign languages are not necessarily bilingual), who might use pronunciations partly inherited from source languages, and “naive” speakers pronouncing loanwords based on the generalizations of Foreign. It is difficult to determine the true nature of the data attested in pronouncing dictionaries. Do the transcriptions of loanwords actually reflect what bilingual speakers realise or are they only representative of “naive” speakers? As it is really difficult to classify or select native speakers according to their abilities in English (if the impact of bilingual speakers on the dictionary-making process turns out to be real), let us review some evidence which will give us some insight on how speakers use the subsystems we have discussed.

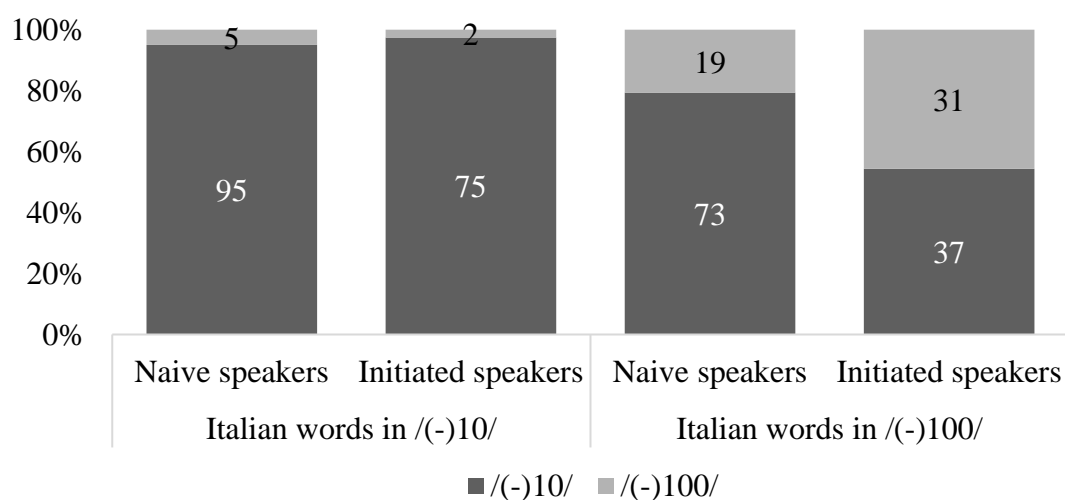
After reporting a dictionary-based study of stress patterns in Italian loanwords in English, P. Fournier (2018) presents a reading experiment with seven native English speakers. The stress match rate in dictionaries is very high between Italian words and Italian loanwords in English. This high stress match rate can be expected to affect the oral productions of native English speakers in the experiment because native English speakers may have intuitions concerning the way Italian words should be stressed. Instead of selecting Italian loanwords, whose pronunciations English speakers might be familiar with, only Italian words which are not attested in English were chosen<sup>23</sup>. Those Italian words were integrated into a set of foreign words, so that it was impossible for the participants to know what was really tested. In order to evaluate the stress match and check the sensitivity of English speakers to the original stress patterns, the tested Italian words submitted to the informants were either stressed on the penultimate or on the antepenultimate syllables. The results showed that words stressed on the penult in Italian were also massively assigned penultimate stress by the participants. This phenomenon was anticipated and cannot be used as evidence for the sensitivity of speakers to the phonological properties of Italian. However, it also appeared that the /(-)100/ stress match rate between Italian and English rises according to the linguistic abilities of English native speakers in Italian. The criterion of exposure to Italian had not been anticipated in the preliminary steps of the study and turned out to be significant while interpreting the results. None of the speakers was bilingual and they had only (for three of them) learnt Italian for one or two years in high school. The results from P. Fournier (2018) are presented in Figure 3.

---

<sup>23</sup> All the selected words have <-VCV#> structures.



Figure 3: Stress patterns produced by the participants in P. Fournier (2018) depending on the stress pattern attested in Italian and on the participants' exposure to Italian



From P. Fournier (2018)

The results show that English native speakers have intuitions concerning the way loanwords should be pronounced, even though this is not a conscious phenomenon:

The oral test, experimented on native English speakers, has led to the following conclusion: English speakers massively stress Italian words on the penultimate syllable whatever the original Italian stress patterns. However, speakers with some knowledge of Italian reproduce more faithfully original Italian stress patterns and the observation of their stressing concerning the Italian items that are stressed in /(-)100/ confirms that the more they grow accustomed, the more they stress in accordance with Italian stress principles. (P. Fournier 2018)

Even if exposure to Italian was not meant to be significant at the beginning of the test, it turned out to be so. But this parameter should be thoroughly analysed, and further research is required to determine the exact effects of exposure to the source language. This preliminary conclusion is also found in Kang (2010): “the closer the contact and the higher the level and rate of bilingualism, the more likely the adapters are to perceive stress contrasts of the input language correctly, in turn leading to the stronger possibility of the preservation of input stress.”

Another experiment is conducted in Fitt (1996), in which the names of 60 European cities were read by young students (aged 13 to 17). The results concerning the stress patterns produced by participants for trisyllabic words are shown in Figure 4.

Figure 4: Stress patterns produced by the participants in Fitt (1996) for trisyllabic city names

	/100/	/010/	/201/
<b>Mostly /010/</b>			
<i>Livorno</i>	1	20	0
<i>Valençay</i>	3	16	0
<i>Bolkesjö</i>	4	10	0
<i>Larisa</i>	1	21	0
<i>Megara</i>	4	18	0
<i>Novoli</i>	1	19	0
<i>Copparo</i>	3	19	0
<i>Ekhnos</i>	7	13	0
<b>Mostly /100/</b>			
<i>Bobbio</i>	20	2	0
<i>Osimo</i>	12	7	0

From Fitt (1996)

The results show that speakers massively use penultimate stress whenever words are felt to be foreign. We assume that these words are associated to §Foreign, for which the default pattern is penultimate stress. Interestingly, *Bobbio* has an ending that is regularly associated to antepenultimate stress (e.g. *áudio*, *fólio*, *pistáchio*, *rádio*, *rátio*), which shows that certain factors may override the default pattern.

Before we conclude this section, let us briefly return to the orthographic subsystems discussed previously. Although there is strong dictionary-based evidence for these subsystems, it is possible that speakers do not use such orthographic subsystems. However, there is psycholinguistic evidence which shows that they are sensitive to these subsystems. Treiman, Kessler and Evans (2007) found that English speakers are more likely to pronounce <c> and <g> as /s/ and /dʒ/, respectively, in nonwords with graphotactic patterns that are typical of Latinate words (e.g. <id>, <ph>, <ous>), and that they are more likely to pronounce these graphemes as /k/ and /g/ in nonwords with graphotactic patterns that are typical of “basic” words. These results suggest that speakers are somehow aware of these orthographic subsystems, a finding which was recently further supported by Treiman, Decker and Kessler (2019). They showed participants pairs of nonwords which either contained consistent or inconsistent graphotactic sequences (e.g. Latinate onset + Latinate ending in *-phalid* vs. “basic” onset + Latinate ending in *-shalid*) and asked them to select the more wordlike item from each pair. They found that participants preferred items with consistent spellings in over 98% of the time, in all three experimental conditions (variation in the onset, in the ending and in medial graphemes). Moreover, let us briefly evoke the case of the three most common “foreign free vowels” shown in Figure 1, i.e. <a> - /ɑ:/, <e> - /eɪ/ and <i> - /i:/ . As argued before, such correspondences can be attributed to a principle of articulatory proximity considering that, in some of the main source languages (e.g. Spanish, Italian, French), the vowels <a>, <e> and <i> are usually pronounced /a/, /e/ and /i/ respectively.

What we have seen in this section allows us to formulate a prediction, which will have to be investigated in future research: when presented with words with identifiable characteristics from one of the subsystems, speakers will tend to use a default pattern which is based on the words available in their lexicon and may diverge from the patterns found in §Core vocabulary. Therefore, first, we expect P. Fournier’s (2018) findings to be replicated: speakers who are familiar with the language from which words were borrowed are expected to diverge

from the default pattern, if the source language contains words which indeed diverge from that pattern, more than speakers with little or no familiarity with the source language. Second, speakers with a larger vocabulary should be more likely to produce such patterns which diverge from §Core patterns than speakers with a smaller vocabulary. These predictions could be tested both through reading and judgment tasks<sup>24</sup>.

#### 4. Relationship with the source languages

This article assumes that English phonology should not be considered as a uniform system, but rather that it is composed of several subsystems. However, another element cannot be overlooked: there is a relationship between source languages and target languages and the borrowing process necessarily implies some phonological interferences between the corresponding source languages and English. One might find parallels between this preliminary assumption and the theoretical principles developed by researchers within the Loanword Phonology framework.

##### 4.1. Loanword Phonology

Researchers working within Loanword Phonology have described two models accounting for the pronunciation of loanwords. Their views are not limited to English and the models are thought to be universal. These views and the debate surrounding them are presented and detailed in Calabrese and Wetzels (2009, 1–3) along with exhaustive references. The first model, called the “phonological stance model” or “the nativization-through-production scenario” (see for example Hyman 1970; Jacobs and Gussenhoven 2005; Paradis and LaCharité 2005), assumes that loanwords are integrated into target languages by bilingual speakers who master the phonological mechanisms of both languages. Loanwords are generally pronounced with the phonological mechanisms ruling the source language, but it appears that the opposite is also possible, i.e. loanwords are pronounced with the phonological properties of the target language. Assimilatory processes are more attested when the principles of the target language are used to pronounce words from other languages. The second model, called the “perceptual stance model” or the “nativization-through-perception scenario” (see among others Silverman 1992; Peperkamp and Dupoux 2008), claims that loanwords are integrated into target languages by speakers who are not familiar with the phonological principles ruling the source language. Loanwords necessarily go through a step of assimilation to the mechanisms of target languages because of the speakers’ inability to perfectly reproduce the pronunciation of source languages. Assimilation is only based on the speakers’ capacity to decode and interpret perceptual cues and reproduce them. Each model relies on a distinct input and it can be concluded that the first model is phonological when the second one is phonetic (Calabrese and Wetzels 2009, 2). One might be tempted to compare the borrowing process with code-switching, with speakers using two distinct linguistic codes in sentences (see for example Myers-Scotton (1993) for a complete description of the mechanisms of code-switching). However, the two phenomena must be distinguished (Haspelmath 2009, 40) because loanwords, once they have been integrated into target languages, truly belong to those target languages. The stylistic dimension that code-switching can refer to is not expressed in the same way in loanwords, whose use is more associated with prestige, but it is true that in the earliest stages of attestation of loanwords in target languages, it can be hard to distinguish loanwords from code-switching.

The two models exposed above echo the arguments found in Davis, Tsujimara and Tu (2012) who summed up the most recent research concerning the pronunciation of loanwords with a taxonomy for loanword prosody. According to them, even if prosodic prominence is

---

<sup>24</sup> Note that local analogical effects should be controlled for. Indeed, Trevian (1993, 246) evokes experiments on nonwords such as *tomino*, which most participants stressed /100/, probably in analogy with *domino*, and contrary to what the rule of “Italian” words (or the default §Foreign stress rule we are assuming) would predict.

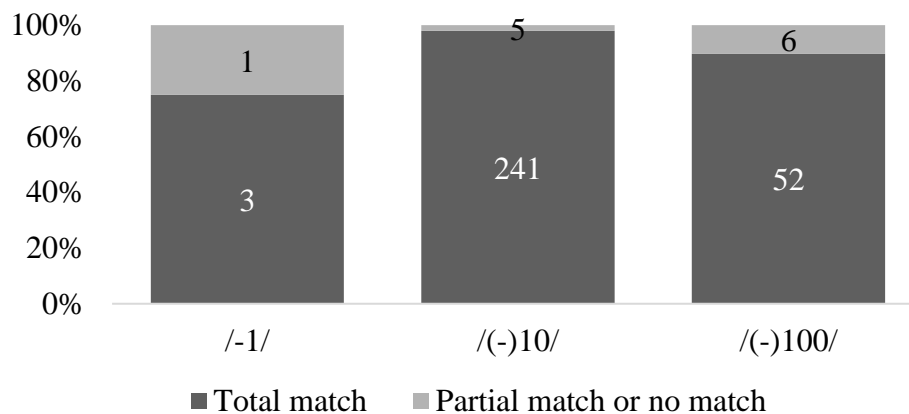
differently represented in each language, there is a necessary phase of adaptation when loanwords are integrated into new languages and two main criteria have to be taken into account and investigated. First, some phonological features of the source language may potentially be reproduced in the target language. Then, if the stressing of loanwords is dealt with using the phonological tools of the target language, it turns out that some target languages may use the general rules which account for the whole lexicon while some others may apply rules which are specific to loanwords.

This last point is of major importance regarding the stressing of loanwords in English. In the previous sections, five pronunciation subsystems operating within the English stress system have been posited and described. We assume that the English subsystems have gained a form of autonomy from the source languages. This means that a word perceived as belonging to §French or §Foreign might be pronounced in ways that diverge from the pronunciation found in the source languages but are consistent with the generalizations of the subsystem.

#### 4.2. Influence of phonological mechanisms from source languages

P. Fournier's analysis of Italian loanwords in English (2018, 2021b) shows that there is an obvious link between some stress phenomena in English and phonological features found in foreign languages. Indeed, these dictionary-based studies show a close-to-perfect stress match between the stress patterns of Italian words and the stress patterns of Italian loanwords in English. In Italian, words are mainly stressed on the penultimate syllable (80%), while antepenultimate stress (16%) and final stress (4%) are less common (Borrelli 2002; Thornton, Jacobini and Burani 1997). The results of the stress match comparison between Italian words and their corresponding Italian loanwords in English found in P. Fournier (2018), as shown in Figure 5, are explicit.

Figure 5: Stress match rate between Italian words and the corresponding loanwords in English



From P. Fournier (2018)

One observes that 296 out of 309 items exhibit identical stress patterns in Italian and in English. Italian words are mainly stressed on the penult and, if only loanwords with penultimate stress are considered (e.g. *affettuoso* → *affettuoso*, *dogána* → *dogána*, *riprésa* → *riprésa*), it is impossible to determine whether penultimate stress in English is the result of the reproduction of the Italian stress pattern or the application of the preferential penultimate stress pattern for §Foreign words in English. However, the survey of other stress patterns shows that the transmission of phonological properties from source languages is a factor that cannot be ignored. Indeed, cases of antepenultimate stress correspondences are also attested (e.g. *cúpola* → *cúpola*, *búfalo* → *búffalo*, *ópera* → *ópera*, *altíssimo* → *altíssimo*, *bróccolo* → *bróccoli*, *enérgico* → *enérgico*, *tímpano* → *tímpani*), as well as final ones (e.g. *tìramisú* → *tìramisú*,

*oimé* → *oimé*, *romanitá* → *Romanitá*, *terribilitá* → *terribilitá*). Stress variation is notably attested mainly with the appearance of penultimate stress, and marginally of antepenultimate stress, in English, which is not attested in Italian (e.g. *omertá* → /010/ ~ /201/; *incógnito* → /2010/ ~ /0100/; *vióla* → /100/ ~ /010/; *furóre* → /010/ ~ /100). This variation phenomenon is probably representative of the conflict between two logics: stress in Italian loanwords is either determined by the parameters ruling the English system (we can mention the rule of “Italian” words and extend it to the preferential stress pattern for loanwords) or by the influence of the original Italian stress pattern. The paradoxical case of <-VIV#> structures (P. Fournier 2021b) is worth mentioning. The rule of “Italian” words only integrates /t, d, n, s, z/ as consonants which are associated to penultimate stress when found in <-VCV#> contexts. The /l/ consonant is excluded because Italian words in <-VIV#> are evenly divided between penultimate or antepenultimate stress. The influence of this structure is not stable enough in Italian. But if we only consider the stress patterns of Italian loanwords in <-VIV#> in English, it turns out that there is a close-to-perfect stress match between Italian words and the corresponding loanwords in English. The results in Figure 6 are taken from P. Fournier (2021b):

Figure 6: Stress match comparison between Italian words and Italian loanwords in English.

<b>Stress match</b> /(-)10/ → /(-)10/	<b>Stress match</b> /(-)100/ → /(-)100/	<b>Variation</b>
7 items <i>cicála</i> → <i>cicála</i> <i>pianóla</i> → <i>pianóla</i> <i>scaglióla</i> → <i>scaglióla</i> <i>raviólo</i> → <i>ravióli</i> <i>biennále</i> → <i>biennále</i> <i>campaníle</i> → <i>campaníle</i> <i>finále</i> → <i>finále</i>	15 items <i>bresáola</i> → <i>bresáola</i> <i>cúpola</i> → <i>cúpola</i> <i>góndola</i> → <i>góndola</i> <i>rúcola</i> → <i>rúcola</i> <i>tarántola</i> → <i>tarántula</i> <i>bróccolo</i> → <i>bróccoli</i> <i>búfalo</i> → <i>búffalo</i>	3 items <i>pérgola</i> → /100/ ~ /010/ <i>tómbola</i> → /010/ ~ /100/ <i>vióla</i> → /100/ ~ /010/

From P. Fournier (2021b)

The variation is extremely limited, and it turns out that the Italian stress pattern is reproduced in English along with the lexical borrowing. Therefore, it means that the influence of phonological principles from source languages prevails here over the application of the rule of “Italian” words operating in §Foreign. The systematic reproduction of Italian stress patterns generates Italian loanwords with penultimate or antepenultimate stress in English. Those antepenultimate stress patterns, which cannot not be accounted for by the English rule of “Italian” words (they could only be considered as exceptions), make it clear that /l/ cannot be integrated into this English stress rule.

Concerning the pronunciation of French loanwords, Castanier (2016, 293–300) reports interesting evolutions in the pronunciation of words ending in <-et> (e.g. *buffet*, *filet*, *trebuchet*). In English, over the past three centuries, the tendency has been stress retraction in disyllables but not in longer words, where some words have even adopted final stress when they had antepenultimate stress at an earlier time (e.g. *clárinet* → *clàrinét*). Castanier points out that some words were re-borrowed with different semantics (e.g. *flageolet* used to refer to a musical instrument but now refers to a type of bean) and that this could explain the evolution of their stress pattern, but that this explanation could not be extended to all the words which have undergone this evolution. He also notes that the pronunciation tends to evolve towards /-ei/, even when the historical pronunciation is /-et/. This parallels the evolution in French from /-et/ to /-e/, but the evolution also extends to words which do not exist in French anymore (e.g.

*marmoset*). In this example, the influence of the source language is somehow indirect, as the way loanwords are pronounced in English is not so much influenced by the pronunciation of individual words as by the fact that the evolution of the graphophonological correspondence <-et> - /-et/ to <-et> - /-e/ has been gradually transmitted to English, so that the new correspondence <-et> - /-ei/ has spread to words which do not exist in contemporary French anymore.

#### 4.3. Limits and autonomy of English subsystems

The influence of phonological parameters from source languages on English is a fundamental phenomenon which cannot be ignored. However, the complexity of this study is to evaluate the exact influence and the limits of phonological interferences from source languages on English. The arguments we have developed in this paper make it clear that we assume that the hypothesis of mere imitations or reproductions cannot account for the whole pronunciation of loanwords in English. The way loanwords enter the English lexicon, possibly through bilingual speakers, shows that the phonological properties from source languages must have had a decisive influence on the pronunciation of borrowings in English. However, we assume that the English phonological system has gained a form of independence with regard to these source languages as well. The stress variation found in loanwords is evidence of conflicts that exist between the influence of the phonological properties of source languages and the autonomation of parameters which are now part of the English system<sup>25</sup>.

If the assumption according to which English phonology is composed of several subsystems is correct, the pronunciation of loanwords could be accounted for by the generalizations of the subsystems they are associated with, even if that leads to mismatches with what is found in the source language. For example, if an Italian loanword whose original Italian stress pattern is /(-)100/ is stressed /(-)10/ in English, we can interpret this as the application of the stress rules of §Foreign. We saw in the previous section that such cases are indeed attested. Therefore, the word is still somehow perceived as “foreign”, but its pronunciation is not directly determined by how it is pronounced in the source language. Instead, it follows the generalizations found in words of the subsystem it belongs to. Although the mismatches might be rare because the generalizations of the subsystems broadly reflect what is found in the source languages, our model predicts that such mismatches will occur when there is a conflict between the pronunciation of the source language and the generalizations found in the subsystem.

Interestingly, there can be stress mismatches between the source language and its pronunciation in English along with the preservation of other “foreign” properties. For example, Bermúdez-Otero (2012) evokes the cases of borrowings from Russian<sup>26</sup>: “Russian *Nínotchka*, *bábushka* > English *Ninótchka*, *babúshka*. It is interesting to note that these loanwords have been nativized metrically, but not segmentally: the clusters [t̪.k] and [ʃ.k] occur tautomorphemically only in borrowings like A[ʃ.k]*kenazi* and A[ʃ.k]*elon*”.

---

<sup>25</sup> Stress variation could also be taken to be a sign of hesitation among English speakers so as to which subsystem a word should be assigned and, as a consequence, hesitation regarding which stress rules should be used.

<sup>26</sup> Such cases are not isolated, although they may not be representative of the whole class of words considered. A search in Upton and Kretzschmar (2017) returned eight trisyllabic words which were borrowed from Russian and have antepenultimate stress in Russian. Five of them are pronounced with penultimate stress in English (*babushka*, *Kerensky*, *Mussorgsky*, *Pavlova* (~ /100/), *Scriabin*) while the remaining three preserve antepenultimate stress (*Horowitz*, *Korsakoff*, *Sholokhov*). Penultimate stress seems to be preserved more often as, out of 19 trisyllabic words in the same dictionary with penultimate stress in Russian, 11 have penultimate stress in English (*Gagarin*, *Gromyko*, *Kalinin*, *Kaluga*, *Lysenko*, *Minkowski*, *Potemkin*, *Odessa*, *Rasputin*, *tovarich*, *tsarevna*), three alternate between /010/ and /100/ (*Kasparov*, *Maksutov*, *Nureyev*) and five are stressed /100/ in English (*Khabarovsk* (~ /201/), *Nabokov*, *Romanov*, *tsarevich*, *Vladimir*).

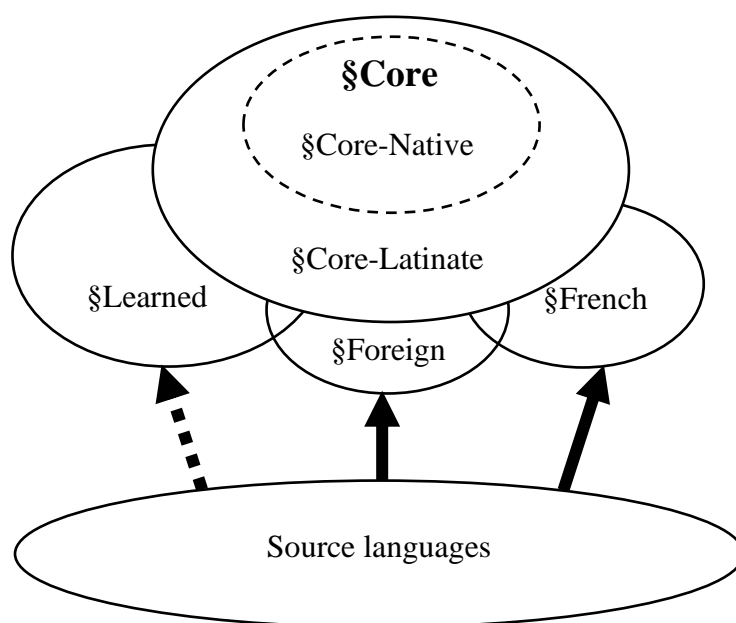
In the analysis we have developed in this paper, the shift to penultimate stress can be argued to be an effect of the affectation of these words to §Foreign, and therefore the application of the preferential penultimate stress pattern, here favoured by the presence of a prefinal consonant cluster. However, in our model, the preservation of the clusters /tʃk/ and /ʃk/ is not a problem as “deviant” phonotactics can be defining properties of the subsystems (as seen for the distribution of /ʒ/ in §French). Therefore, such cases show us that the integration of foreign words into one of the subsystems in English may not affect all of their phonological properties in the same way. For words which can be associated to §Foreign, we expect to see more stress shifts to the penult if stress is on the final or antepenultimate syllable in the source language, as this pattern conflicts with the default pattern in §Foreign. However, we assume that unusual phonotactics are more likely to be preserved as they do not conflict with properties of the subsystems in English.

### 5. Possible formalism

As we pointed out previously, we are here exploring a research hypothesis, and little empirical work seeking to specifically test out this hypothesis has yet been conducted. Consequently, it would be premature to seek to propose a definite theoretical account of empirical facts which are still to be refined and clarified. However, in this section, we briefly explore what existing theoretical tools are available and how they might be used to account for the system we are assuming.

First, let us sketch out a schematic representation of the system we are positing. This representation is shown in Figure 7.

Figure 7: Schematic representation of the subsystems assumed in this paper

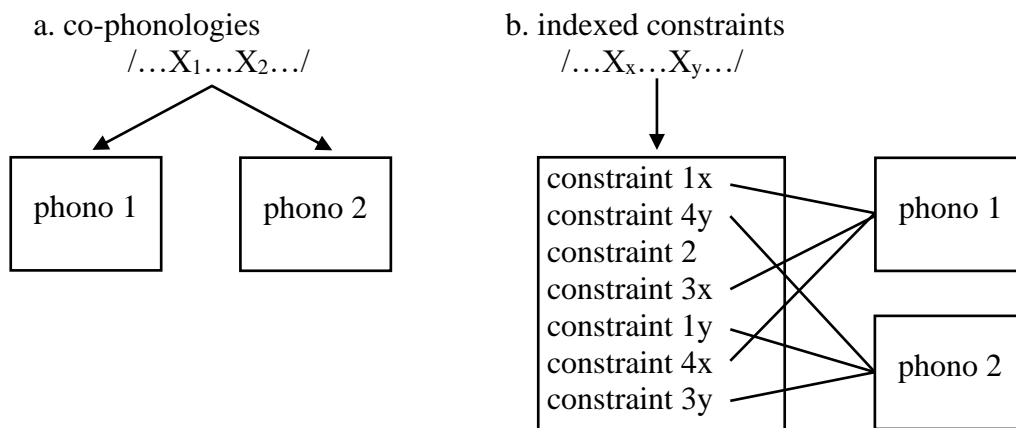


Let us briefly comment this representation. As was seen in the previous sections, we assume that the two components of §Core are strongly intertwined and therefore the boundary between these two systems is not as strong as it is for the other subsystems. The subsystems §Learned, §Foreign and §French are partly merged into §Core, as not all the words, or not all properties, are systematically distinct from those of §Core (e.g. *indigo* could be analysed as §Foreign but has antepenultimate stress, which is expected in §Core yet not in §Foreign, and all spelling-to-sound correspondences are the regular correspondences of §Core). Therefore,

we assume that, when there is no specificity in a particular domain (e.g. stress, morphology, etc.), words will follow the most general principles, i.e. those of §Core. Figure 3 also shows an overlap between §Learned and §Foreign, and between §Foreign and §French. This is because we saw that certain properties are shared by these subsystems (e.g. the three main foreign free vowels for §Foreign and §French). Finally, the arrows represent the influence of source languages on the subsystems in English. We have seen that some of the properties of the subsystems are directly imported from the source languages (e.g. final stress in It. *romanità* → Eng. *Romanità*), although we have argued that the subsystems have a form of autonomy from these source languages. We chose to represent the arrow leading to §Learned with a dashed line to distinguish it from the other two subsystems. Indeed, the source languages for §Foreign and §French are still “living” languages and the influence can be significantly stronger and more direct than it can be in the case of §Learned, for which the source languages, Greek and Latin, are “dead” languages.

In the phonological literature, two main tools have been used for lexically-specific patterns such as those found in loanwords: co-phonologies (Anttila 2002; Bermúdez-Otero and McMahon 2006; Itô and Mester 1995b) and lexically-indexed constraints (Pater 2000, 2009). These two systems can be represented as in Figure 8.

Figure 8: Representations of co-phonologies and indexed constraints



From Scheer (2011, §482)

Co-phonologies consist in assigning different parts of the lexicon to different phonological grammars, although these grammars may share certain properties, while indexed constraints consist in using a single Optimality Theory (Prince and Smolensky 1993) constraint ranking for the whole lexicon, but assigning certain constraints only to certain parts of the lexicon. As Scheer (2011, §482) points out, “the difference is not easy to transform into contrasting predictions”. It seems to us that a system such as that used for Japanese by Itô and Mester (1995a), which includes several constraints which overlap over different parts of the lexicon, is particularly promising to account for the system we have sketched out in this paper.

The system as we understand it is quite complex in that the subsystems are not exactly waterproof as they may overlap and share certain properties and that words sometimes appear not to display all of the properties of a given subsystem (e.g. the words in *-age* discussed by Castanier (2016, 225), which show stress retraction while keeping typically §French graphophonological correspondences). It is possible that the generalizations which we have identified as specificities of certain subsystems are probabilistic and apply more or less efficiently across all subsystems. In that regard, the whole system would probably be better



formalized using probabilistic models such as Max-Ent-OT (Goldwater and Johnson 2003). Therefore, although it seems too early to propose a formal account of our proposal, we have seen that a number of existing tools in phonological theory should allow us to capture the system we are assuming for English.

## Conclusion

To sum up, we are proposing that English should be analysed as a series of imbricated subsystems, which correspond to certain lexical classes of words and formatives borrowed from other languages. Although these subsystems are influenced by the properties of the source languages, they have a form of autonomy, in the sense that the generalizations that regulate the formatives which belong to a particular system do not necessarily correspond to those found in the corresponding source language. We have sought to review as much of the available evidence for these subsystems as we are aware of. However, this review should not be seen as exhaustive, and future research will certainly refine our proposal. We have seen that there is some evidence that speakers are indeed sensitive to these subsystems and we have proposed some predictions which could be tested to further support our proposal. We have focused on the available empirical evidence and have not sought to provide a formal account of the overall grammar we are proposing, but we have seen that phonological theory already possesses theoretical tools which should allow us to formally account for the system we are proposing. Finally, one question which we have not dealt with in this paper and could complicate a comprehensive formal account is that of gradual nativization, i.e. the fact that loanwords present in the subsystems may gradually be integrated into §Core. This means that, at a given point in time, there might be words with certain, but not all, properties of a given subsystem. As discussed earlier, this is something which can be found, for example, for words in *-age*, which tend to adopt the §French (grapho)phonological shape /-ɑ:(d)ʒ/ although they do not have final stress. Therefore, it remains to be seen whether a core-periphery structure of the type proposed by Itô & Mester (1995b) for Japanese is necessary on top of the proposed subsystems.

## References

- Albrow, Kenneth H. 1972. *The English Writing System: Notes towards a Description*. London: Longman.
- Anttila, Arto. 2002. “Morphologically Conditioned Phonological Alternations.” *Natural Language and Linguistic Theory* 20 (1): 1–42.
- Arndt-Lappe, Sabine. 2014. “Synchronic and Diachronic Analogy in Suffix Rivalry: The Case of English *-ity* and *-ness*.” *English Language and Linguistics* 18 (2), 497–548. doi: 10.1017/S136067431400015X.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Bauer, Laurie, Rochelle Lieber, and Ingo Plag. 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.
- Bermúdez-Otero, Ricardo. 2012. “The Architecture of Grammar and the Division of Labour in Exponence.” In *The Morphology and Phonology of Exponence*, edited by Jochen Trommer, 8–83. Oxford: Oxford University Press.
- . 2018. “Stratal Phonology.” In *The Routledge Handbook of Phonological Theory*, edited by S.J. Hannahs and Anna R. K. Bosch, 100–134. Abingdon: Routledge.
- Bermúdez-Otero, Ricardo, and April McMahon. 2006. “English Phonology and Morphology.” In *The Handbook of English Linguistics*, edited by Jochen Trommer, 382–410. Oxford: Oxford University Press.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt and Company.
- Borrelli, Doris. 2002. *Raddoppiamento Sintattico in Italian. A Synchronic and Diachronic Cross-Dialectal Study*. New York & London: Routledge.

- Burzio, Luigi. 1994. *Principles of English Stress*. New York: Cambridge University Press.
- Calabrese, Andrea, and W. Leo Wetzels. 2009. *Loan Phonology*. Amsterdam and Philadelphia: John Benjamins Publishing.
- Carney, Edward. 1994. *A Survey of English Spelling*. London: Routledge.
- Castanier, Jérémy. 2016. *L'évolution accentuelle du lexique anglais contemporain appréhendée à travers les dictionnaires de prononciation (XVIII<sup>e</sup>-XXI<sup>e</sup> siècles)* [unpublished doctoral dissertation]. Université de Poitiers.
- Chadelat, Jean-Marc. 2000. *Valeur et fonctions des mots français en anglais à l'époque contemporaine*. Paris et Montréal: L'Harmattan.
- Chirol, Laure. 1976. *Les "mots français" et le mythe de la France en anglais contemporain*. Paris: Klincksieck.
- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Church, Kenneth. 1985. "Stress Assignment in Letter-to-Sound Rules for Speech Synthesis." *The Journal of the Acoustical Society of America* 78 (S1).
- Coates, Richard. 1999. *Word Structure*. New York: Routledge.
- Collie, Sarah. 2007. *English Stress Preservation and Stratal Optimality Theory* [unpublished doctoral dissertation]. University of Edinburgh.
- Dabouis, Quentin. 2016a. "Is the Adjectival Suffix *-al* a Strong Suffix?" *Anglophonia* 21.
- . 2016b. *L'accent secondaire en anglais britannique contemporain* [unpublished doctoral dissertation]. Université de Tours.
- . 2017. "Semantically Opaque Prefixes and English Phonology." Paper presented at the 14<sup>th</sup> Old World Conference in Phonology, 20-22<sup>th</sup> February, Düsseldorf.
- . 2018. "The Pronunciation of Vowels with Secondary Stress." *Corela* 16 (2).
- . Forthcoming "English Phonology and the Literate Speaker: Some Implications for Lexical Stress." In *English Word Stress: Theories, Data and Variation*, edited by Nicolas Ballier, Jean-Michel Fournier, Anne Przewozny and Eiji Yamada.
- Dabouis, Quentin, Jean-Michel Fournier, Pierre Fournier, and Marjolaine Martin. Forthcoming. "English Word Stress and the Guierrian School." In *English Word Stress: Theories, Data and Variation*, edited by Nicolas Ballier, Jean-Michel Fournier, Anne Przewozny and Eiji Yamada.
- Davis, Stuart, Natsuko Tsujimura, and Jung Yueh Tu. 2012. "Toward a Taxonomy of Loanword Prosody 11." *Catalan Journal of Linguistics*.
- Deschamps, Alain. 1994. *De l'écrit à l'oral et de l'oral à l'écrit*. Paris: Ophrys.
- Dresher, Elan B. 2013. "The Influence of Loanwords on Norwegian and English Stress." *Nordlyd* 40 (1): 55–65.
- Fidelholtz, James L. 1975. "Word Frequency and Vowel Reduction in English." *Chicago Linguistic Society* 11: 200–213.
- Fitt, Susan. 1996. "Syllable Weight, Stress Patterns and Nativisation." In *Proceedings of the Edinburgh Linguistics Department Conference*, 40–51. University of Edinburgh.
- Fournier, Jean-Michel. 2007. "From a Latin Syllable-Driven Stress System to a Romance versus Germanic Morphology-Driven Dynamics: In Honour of Lionel Guierre." *Language Sciences* 29 (2–3): 218–36. doi:10.1016/j.langsci.2006.12.016.
- . 2010a. "Accentuation lexicale et poids syllabique en anglais : l'analyse erronée de Chomsky et Halle." Paper presented at the 8<sup>th</sup> meeting of the French Phonology Network (RFP), 1<sup>st</sup>–3<sup>rd</sup> July, Université d'Orléans.
- . 2010b. *Manuel d'anglais oral*. Paris: Ophrys.
- Fournier, Pierre. 2013. "Accentuation des composés savants en anglais contemporain : identification morphosémantique et enjeu phonologique." *CORELA* 11 (2).

- . 2016a. “L’accentuation des emprunts français dans les dictionnaires anglais de prononciation.” In *Proceedings of the XVII EURALEX International Congress - Lexicography and Linguistic Diversity*, edited by Tinatin Margalitadze and George Meladze, 424–35. Tbilisi: Ivane Javakhishvili Tbilisi State University.
- . 2016b. “Nouvelles perspectives sur l’accentuation des emprunts en anglais contemporain.” *Itinéraires* 2015–2.
- . 2017. “Vers une tentative d’uniformisation du traitement accentuel des composés savants suffixés en anglais contemporain.” *Anglophonia* 23.
- . 2018. “Stress Assignment in Italian Loanwords in English and Its Impact on the Stressing of Foreign Words by Native English Speakers.” *CORELA* 24.
- . 2021a. “Les emprunts au gaélique écossais en anglais.” In *L’Écosse: La Différence*, edited by Sabrina Juillet-Garzon, Pierre Fournier and Arnaud Fiasson. Presses Universitaires de Franche-Comté.
- . 2021b. “Transcription et formalisation de l’accentuation des mots ‘italiens’ en anglais contemporain.” In *Transcrire, Écrire, Formaliser 2*, edited by Fabienne Toupin and Catherine Collin, 53–68. Presses Universitaires de Rennes.
- Giegerich, Heinz J. 1999. *Lexical Strata in English: Morphological Causes, Phonological Effects*. Cambridge: Cambridge University Press.
- Goldwater, Sharon, and Mark Johnson. 2003. “Learning OT Constraint Rankings Using a Maximum Entropy Model.” *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–20.
- Guierre, Lionel. 1979. *Essai sur l’accentuation en anglais contemporain : éléments pour une synthèse* [unpublished doctoral dissertation]. Université de Paris-VII.
- . 1988. “Les orphelins de la suffixation.” In *4<sup>e</sup> colloque d’avril sur l’anglais oral*, 147–52. Villetaneuse: Université de Paris-Nord, CELDA, diffusion APLV.
- Hammond, Michael. 1999. *The Phonology of English: A Prosodic Optimality-Theoretic Approach*. Oxford: Oxford University Press.
- Haspelmath, Martin. 2009. “Lexical Borrowing: Concepts and Issues.” In *Loanwords in the World’s Language: A Comparative Handbook*, edited by Martin Haspelmath and Uri Tadmor, 35–54. New York and Berlin: de Gruyter Mouton.
- Hay, Jennifer. 2003. *Causes and Consequences of Word Structure*. London: Routledge.
- Hay, Jennifer, and Harald Baayen. 2003. “Phonotactics, Parsing and Productivity.” *Italian Journal of Linguistics* 15 (1): 99–130.
- Hayes, Bruce. 1982. “Extrametricity and English Stress.” *Linguistic Inquiry* 13 (2): 227–76.
- Heuven, Walter V. J. Van, Pawel Mandera, Emmanuel Keuleers, and Marc Brysbaert. 2014. “Subtlex-UK: A New and Improved Word Frequency Database for British English.” *Quarterly Journal of Experimental Psychology* 67: 1176–90.
- Hyman, Larry. 1970. “The Role of Borrowing in the Justification of Phonological Grammars.” *Studies in African Linguistics* 1 (1): 1–48.
- . 1975. *Phonology: Theory and Analysis*. New York: Holt, Rinehart and Winston.
- Itô, Junko, and Armin Mester. 1995a. “Japanese Phonology.” In *The Handbook of Phonological Theory*, edited by John Goldsmith, 816–38. Oxford: Blackwell.
- . 1995b. “The Core-Periphery Structure of the Lexicon and Constraints on Reranking.” In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, edited by Jill Beckman, Laura Dickey and Suzanne Urbanczyk, 181–209. Amherst, MA: GLSA.
- . 1999. “The Phonological Lexicon.” *The Handbook of Japanese Linguistics*, edited by Natsuko Tsujimura, 62–100. Malden, MA and Oxford, UK: Blackwell Publishers.
- Jacobs, Haike, and Carlos Gussenhoven. 2005. “Loan Phonology: Saliency, the Lexicon and Optimality Theory.” *Linguistic Inquiry* 36 (2): 223–58.

- Kang, Yoonjung. 2010. "Tutorial Overview: Suprasegmental Adaptation in Loanwords." *Lingua* 120 (9): 2295–2310.
- Kaye, Jonathan. 1995. "Derivations and Interfaces." *Frontiers of Phonology*, edited by Jacques Durand and Francis Katamba, 289–332. London and New York: Longman.
- Martin, Marjolaine. 2011. *De l'accentuation lexicale en anglais australien standard contemporain* [unpublished doctoral dissertation]. Université de Tours.
- Minkova, Donka. 1997. "Constraint Ranking in Middle English Stress-Shifting." *English Language and Linguistics* 1 (1): 135–75.
- . 2006. "Old and Middle English Prosody." In *The Handbook of the History of English*, edited by Ans van Kemenade and Bettelou Los, 95–125. Oxford: Blackwell Publishing.
- Montgomery, Scott. 2001. *The Case for Synchronic Orthographic Primacy: The Effect of Literacy on Phonological Processing* [unpublished doctoral dissertation]. University of Edinburgh.
- Myers-Scotton, Carol. 1993. "Common and Uncommon Ground: Social and Structural Factors in Codeswitching." *Language in Society* 22 (4): 475–503.
- Paradis, Carole, and Darlene LaCharité. 2005. "Category Preservation and Proximity versus Phonetic Approximation in Loanword Adaptation." *Linguistic Inquiry* 36 (2): 223–58.
- Pater, Joe. 1994. "Against the Underlying Specification of an Exceptional English Stress Pattern." *Toronto Working Papers in Linguistics* 13.
- . 2000. "Non-Uniformity in English Secondary Stress: The Role of Ranked and Lexically Specific Constraints." *Phonology* 17: 237–74.
- . 2005. "Learning a Stratified Grammar." *Proceedings of the 29th Boston University Conference on Language Development*, no. 2004: 482–92.
- . 2009. "Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution." In *Phonological Argumentation: Essays on Evidence and Motivation*, edited by Steve Parker, 123–54. London: Equinox.
- Peperkamp, Sharon, and Emmanuel Dupoux. 2008. "A Typological Study of Stress 'Deafness.'" *Laboratory Phonology* 7: 203–40.
- Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter.
- . 2002. "The Role of Selectional Restrictions, Phonotactics and Parsing in Constraining Suffix Ordering in English." In *Yearbook of Morphology 2001*, edited by Geert Booij and Jaap Marle, 285–314. Dordrecht: Kluwer Academic Publishers.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar* [unpublished research]. Rutgers University and University of Colorado.
- Raffelsiefen, Renate. 1993. "Relating Words: A Model of Base Recognition. Part I." *Linguistic Analysis* 23: 3–161.
- . 2015. "Phonological Restrictions on English Word-Formation." In *Word-Formation: An International Handbook of the Languages of Europe*, edited by Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer, 894–917. Berlin and Boston: Mouton de Gruyter.
- Ryan, Des. 2017. *Principles of English Spelling Formation* [unpublished doctoral dissertation]. Trinity College Dublin.
- Scheer, Tobias. 2011. *A Guide to Morphosyntax-Phonology Interface Theories. How Extra-Phonological Information Is Treated in Phonology since Trubetzkoy's Grenzsingnale*. Berlin: Mouton de Gruyter.
- Silverman, Daniel. 1992. "Multiple Scansions in Loanword Phonology: Evidence from Cantonese." *Phonology* 9 (2): 289–328.

- Tadmor, Uri. 2009. “Loanwords in the World’s Languages: Findings and Results.” In *Loanwords in the World’s Languages: A Comparative Handbook*, edited by Martin Haspelmath and Uri Tadmor, 55–75. New York and Berlin: de Gruyter Mouton.
- Thornton, Anna M., Claudio Jacobini, and Cristina Burani. 1997. “BDVDB -Una base di dati sul vocabolario di base della lingua italiana.” Roma: Bulzoni.
- Tournier, Jean. 1985. *Introduction descriptive à la lexicogénétique de l’anglais contemporain*. Paris et Genève: Champion - Slatkine.
- . 1993. *Précis de lexicologie anglaise*. 3<sup>rd</sup> edition. Paris: Nathan.
- Treiman, Rebecca, Kristina Decker, and Brett Kessler. 2019. “Adults’ Sensitivity to Graphotactic Differences within the English Vocabulary.” *Applied Psycholinguistics* 40: 167–82.
- Treiman, Rebecca, Brett Kessler, and Rochelle Evans. 2007. “Anticipatory Conditioning of Spelling-to-Sound Translation.” *Journal of Memory and Language* 56: 47–53.
- Trevian, Ives. 1993. *Phonographématique, phonologie et morphophonologie des consonnes en anglais contemporain* [unpublished doctoral dissertation]. Université de Paris VII.
- . 2003. *Morphoaccentologie et processus d’affixation de l’anglais*. Bern: Peter Lang.
- . 2007. “Stress-Neutral Endings in Contemporary British English: An Updated Overview.” *Language Sciences* 29 (2–3): 426–50.
- . 2015. *English Suffixes: Stress-Assignment Properties, Productivity, Selection and Combinatorial Processes*. Bern: Peter Lang.
- Upton, Clive, and William A. Kretschmar. 2017. *The Routledge Dictionary of Pronunciation for Current English*. Abingdon and New York: Routledge.
- Wells, J.C. (1992) 2008. *Longman Pronunciation Dictionary*. 3<sup>rd</sup> edition. London: Longman.