5 Dictionaries, newspaper archives and corpora

5.1 Dictionaries: The quantitative dimension at work

Over the past decades, the assimilation of Anglicisms into the Italian language has continuously raised contrasting opinions and attitudes among linguists. The opponents of exogenous influences compare the input of Anglicisms to a 'flood' or an 'invasion' that will distort the identity of Italian. Others downplay the dangers of the Anglicization of Italian, arguing that it is a marginal phenomenon, largely limited to low-frequency, specialized vocabulary. Counting how many Anglicisms are part of the Italian lexicon is a question that many linguists have tried to answer (Klajn 1972; Cartago 1994; Serafini 2001; Bistarelli 2008; De Mauro 2005; Furiassi 2008; Gualdo and Scarpino 2007; Pulcini 2007, 2017). In this chapter, the numerical data provided by some existing dictionaries will be used as a starting point for the quantification of Anglicisms that have been imported and assimilated into Italian and which, as will be shown, can only be indicative but not definitive. Before delving into the analysis of Anglicisms in dictionaries, the composition of the Italian lexicon will be examined in order to get an objective perspective on the share of Italian vocabulary that we are focusing on.

De Mauro (2005) explains that the lexicon of any language can be grouped into three broad categories: a. heritage words; b. exogenous words, borrowed from other languages (non-adapted, adapted and calques); c. endogenous (new) formations, created according to the word formation processes characteristic of a language, especially affixation and composition, starting from lexical bases of groups a. and b.[74] De Mauro (2005: 210) showed that 94.8% of the Italian lexicon is constituted by traditional (heritage) lexis and subsequent formations still in use (categories a. and c.), while the remaining 5% includes words taken from exogenous sources (called 'exoticisms'), including Greek (1.4%) and Latin borrowings (0.8%), and borrowings from modern languages (3%), in the form of both adapted (1.6%) and non-adapted (1.4%) borrowings. Moreover, research into samples of texts which typically contain many foreign words (e.g. texts about sport) proved that the frequency of foreign borrowings amounted to 0.48%. These data justify

---

[74] De Mauro explains (2005: 127) that in the English language this tripartite division is clear-cut: 10% of heritage vocabulary from the Anglo-Saxon stock; 76.5% of borrowings from other languages (42% French, 25% Latin, 4% Spanish, 4% Italian, 1.5% other languages, which makes English the most 'Latinized' and 'Romance' language of the non-Latin languages); 13.5% of endogenous formations, based on heritage or borrowed words. In Romance languages like Italian this distinction is not easy to identify, as both the heritage vocabulary and endogenous formations are Latin-based.

De Mauro's favourable attitude to foreign borrowings, given that, according to his data, their number and frequency rate are low and Italian speakers seem to favour heritage lexis to foreign words. According to De Mauro, foreign words are naturally preferred in advertising and in texts where the foregrounding power of the exotic word is exploited for stylistic reasons, but not in the general language of everyday use, both spoken and written. In addition, he shows that nearly half of the terms marked as 'exoticisms' [ES] in the GDU (De Mauro 2007) are also marked as 'technical and specialized' [TS], that is, they belong to the periphery of the lexicon (see footnote 75 on the layers of the Italian lexicon).

The next important step is to determine the share of English borrowings within the total number of exogenous borrowings. On the basis of the lemmas recorded by De Mauro and Mancini (2003), which is based on the GDU, the number of words originating from foreign languages amounts to 11,104 (4.2%), including 8,196 (3.1%) which are from English. Yet, by focusing on lemmas marked as [ES] (exoticism), i.e., non-adapted borrowings from English, the number goes down to 5,510 (2.1%). Considering the total number of non-adapted borrowings from all languages, it follows that the large majority of all foreign borrowings are from English (73.8%), which makes English the strongest donor language to Italian, albeit not the only one. To sum up, the incidence of Anglicisms in the whole word list of the GDU is 3.1% (of which 2.1% are non-adapted and 1% are adapted).

Similar numerical evaluations have been carried out by other authors, with percentages that are not greatly divergent (Pulcini 2007). Gualdo and Scarpino (2007) provide the same percentage of the input of Anglicisms (3.1%) based on the lemma list of the dictionary *Il Sabatini Coletti* (Sabatini and Coletti 2004). According to older data examined by Klajn (1972), the number of direct Anglicisms discussed in his book is 2,150 (of which 1,600 are non-adapted and 550 are adapted), with an overall incidence of Anglicisms in the whole Italian lexicon amounting to 1.7%. In short, despite numerical differences, there is a consensus among linguists that the share of foreign words attested in Italian and the number of Anglicisms are relatively small.

A different impression of the numerical impact of Anglicisms on Italian is given by their increase over time. Figure 5.1 shows the growth in the number of Anglicisms (non-adapted and adapted) in the GDU (2007) and *Nuovo Devoto-Oli 2022*, which has been remarkable in both dictionaries, even though they differ in size, with the GDU containing 260,709 entries and *Nuovo Devoto-Oli 2022* (digital edition) 110,000 entries. The GDU covers the years up to 2006, whereas *Nuovo Devoto-Oli 2022* is updated every year. What is particularly striking is the rise in the number of Anglicisms in the post-unification years (1851–1900 in Figure 5.1), which is higher than the number of Anglicisms imported in all previous centuries. This makes this period a turning point for the development of the Italian national

language. The increase in the 20th century is even more marked, especially in the second half and across the new millennium. Accordingly, these data may justify the alarms raised by some language observers. On the other hand, data need to be interpreted not only in quantitative terms but also in qualitative ones. Following De Mauro's stance on this matter, it is necessary to weigh these numbers against other decisive linguistic factors, such as their sociolinguistic distribution (their currency and usage frequency in the general language or in specialized domains), and their lifetime. In fact, many words may enter a language for a short period and then fall out of use; since many dictionaries are still conservative reference tools, Anglicisms may secure a place in a dictionary and then firmly keep it for a long time, possibly forever, even when their use is extremely limited (cf. 6.5 on obsolescence).[75]
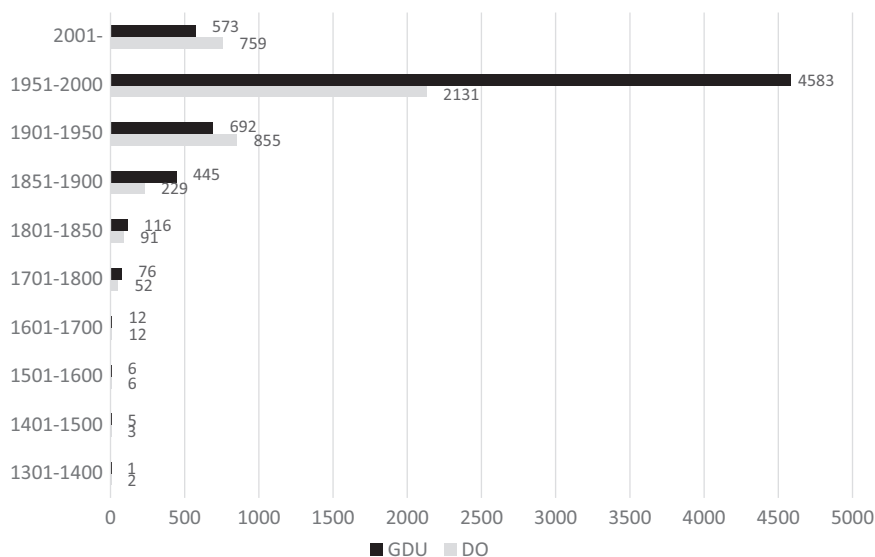

Figure 5.1: Number of Anglicisms in the GDU (2007) and *Nuovo Devoto-Oli 2022.*

---

75 De Mauro and Ferreri (2005) provide a wealth of statistical data, based on their own calculations carried out on *Grande dizionario italiano dell'uso* (2007) and studies carried out by other Italian linguists. They show that great variation on the composition of the Italian lexicon can be found depending on which share one focuses on. The Italian fundamental vocabulary (*fondamentali*) includes c. 2,000 words that occur in 90% of all texts. Among the fundamental vocabulary, the only Anglicisms included are *bar*, *film* and *sport*. Another layer of 3,000 words belong to the category of high use ([AU] *alto uso*) (words understood by any user with average education) and 2,300 high availability words ([AD] *alta disponibilità*) (denoting objects, facts and experiences

As shown in Figure 5.1, the number of Anglicisms recorded in the two dictionaries cited above varies greatly, owing to, in the first place, their different sizes, and to other characteristics such as the time frame, their macrostructure and criteria of inclusion of different types of Anglicisms (Pulcini 2008a, 2010). Nevertheless, the numerical data provide a good picture of the strong upward trend of the input of Anglicisms in Italian since the mid-20th century.

The retrieval of Anglicisms for the creation of a new lexicographic tool with its own specific characteristics generally starts from already existing dictionaries. For the creations of GLAD, some reference dictionaries[76] were identified as suitable sources for the collection of candidate Anglicisms. The reasons for this choice lies in the fact that two of these dictionaries are specifically focussed on Anglicisms, namely the *Dizionario degli Anglicismi nell'italiano postunitario* by Rando (1987) and the *Dictionary of European Anglicisms* by Görlach (2001). One of the chosen dictionaries is devoted to foreign borrowings in the Italian language, namely *Parole straniere nella lingua italiana* by De Mauro and Mancini (2003). Finally, two general dictionaries of Italian, available in digital format, yearly updated and rich in neologisms, namely *Nuovo Devoto-Oli 2022* and *Zingarelli 2022* have been systematically consulted. In the following analysis, the chosen dictionaries will be presented in chronological order and their macrostructural characteristics will be briefly described, with the entry for *jeep* as a sample illustration. A comparative analysis of the letter J entries recorded by these dictionaries, illustrated in Section 5.2, will show the diversity in the Anglicisms recorded in these sources and the use of the comparative method adopted for the selection of GLAD's entries.

– *Dizionario degli Anglicismi nell'italiano postunitario* (1987)
The *Dizionario degli Anglicismi nell'italiano postunitario* by Rando (1987) is the first and only paper dictionary of Anglicisms available for Italian. As declared in the preface to the dictionary, the period considered is from the post-unification time (Italy was unified in 1861), when a decline in lexical borrowing from French gave way to an increase of borrowings from English. The actual number of entries is 2,500 but many forms are embedded within the same headword. The

---

that are known by all speakers). The share of 'common vocabulary' includes 45,000 items, which are known by speakers with a medium-high level of education. Other layers of vocabulary are 'technical and scientific terms' [TS], 'regionalisms' [RE], 'low-use words' ([BU *basso uso*), 'obsolete' [OB] and 'foreign words' ([ES] 'exoticism'). It follows that any evaluation on the impact of English (or any other foreign language) must be measured against the layer of the lexicon in which borrowings have penetrated.

76 The lexicographic record of Anglicisms in Italian began towards the end of the 19th century (Fanfani and Arlia 1877; Panzini 1905; Monelli 1943). The purist intentions of these early contributions were discussed in chapter 2 (cf. 2.4).

author states that the total number of Anglicisms recorded is approximately 4,200, including 3,100 direct loanwords (1,503 non-adapted and 1,600 adapted Anglicisms) and calques (1,100). The types of borrowings included in Rando's dictionary reflect his open definition of an Anglicism:

> Per anglicismo (specificatamente *anglicismo italiano*) si intende ogni vocabolo o accezione di vocabolo che sia di origine inglese o che sia giunto nella nostra lingua tramite l'inglese nelle varietà britannica e americana ma anche in quelle dell'Australia, del Canadà e del Sud Africa. (Rando 1987: xvi)

> [By Anglicism (specifically an *Italian Anglicism*) is meant any word or word sense which is of English origin or has arrived in our language through English in its British and American varieties but also in those of Australia, Canada and South Africa.]

As stated by the author himself, the sources for this dictionary are books, magazines and newspapers, but also spoken sources like radio, television and cinema. The most common fields represented in this dictionary are sport, games and entertainment, science and technology, tourism, business and commerce, politics and warfare. Rando discusses the lexicographic problem of identifying indirect Anglicisms when their form is similar to Italian words (e.g. *editoriale*)–a problem discussed at length in this volume, see Section 4.1–and argues that both linguistic and socio-historical information should guide the lexicographer for the selection of Anglicisms. The result of this pioneering dictionary is a collection of non-adapted and adapted Anglicisms, calques, internationalisms, pseudo-Anglicisms, abbreviations and phrases, many of which are common and recognizable Anglicisms, while others appear extremely infrequent, such as names of animals (e.g. *petrello* from *petrel*, sea bird, better known in Italian as *procellaria*; *vonga-vonga* from Australian bird *wonga-wonga*), highly technical terms (*sincrotrone* from *syncrotron*), and many highly specific acronyms (e.g. JATO, Jet Assisted Take-Off). What also strikes the user is the inclusion of items of encyclopedic nature, like proper names of institutions (*Casa Bianca* from *White House, Associated Press, Lions Club*), trademarks not turned into generic nouns (*Clarks*, *Gillette*, *Rolls Royce*), popular and dialectal variants (*casimira* for cashmere; *piccio* for picture) and many scientific terms of international circulation (e.g. *pennicillina* and *antibiotico*).[77] Some entries are questionable and awkward, such as *credere in* (from

---

77 As discussed in 4.4 many scientific terms can be classified as internationalisms, being made up of neo-classical elements and being attested in many European languages with similar forms and meanings. We argue that basing the attribution of the status of Anglicisms on the nationality of the inventors is a weak criterion. In the examples quoted here, *pennicillina* was discovered by the Scottish scientist Alexander Fleming, and the term *antibiotico* was coined by the biochemist Selman Abraham Waksman, born in Ukraine and naturalized as an American citizen.

to believe in), *perdere l'autobus* (from *to miss the bus*), *Rome by night*, *segnorina* (prostitute), and *tisi galoppante* (galloping tuberculosis). The unbalanced selection of this dictionary's word list is also discussed by Fanfani in his review (1991). Besides the criticism addressed to the inclusion of the items listed above, Fanfani points out that this dictionary presents many inaccuracies in the dates of adoption, the inclusion of Gallicisms and Germanisms, as well as words of dubious English provenance, casuals and quotation words, and dialectal forms featuring in non-standard varieties of Italian. A second edition with corrections was never published, so that the *Dizionario degli Anglicismi nell'italiano postunitario*, despite its weaknesses, remains an important reference work for the description of the lexical influence of English on Italian.

Figure 5.2 shows the entry of the Anglicism *jeep*. The microstructure of the entry indicates grammatical class and gender (s.f., substantive, feminine), followed by a definition, the indication of the plural form (invariable or with -*s*). The grammatical information is followed by possible adaptations: in this case, *gippe* in the Tuscan dialect and the rare form *jep* appear irrelevant to standard Italian. Derivatives and compounds are inserted at the end of each entry. In this case, the

jeep, *s.f.* (in origine, 1944-46, anche *s.m.*). Camionetta scoperta a 4 o 5 posti, molto robusta e adatta alla marcia su terreni difficili. Plurale: **jeep** (talvolta **jeeps**). Adattamenti: **geep,** *s.f.;* **gip** *s.m.* (a 1945); **gippe** (nel toscano); **jep** *s.f.* (raro). Deriv.: **gippone,** *s.m.* ' camionetta della polizia ' Amer. *jeep* trascrizione secondo le regole fonetiche della sigla *G.P.* (*general purpose* ' uso generale '); l'autoveicolo, ideato negli Stati Uniti verso il 1936, venne ampiamente usato dagli eserciti alleati durante la II guerra mondiale e fu introdotto in Italia con l'occupazione passando, poi, anche ad uso civile. **Jeep-girl,** *s.f.,* voce dell'occupazione: ragazza che andava volentieri in jeep son soldati alleati; adattata con **gip(p)ista,** *s.f.* ⟨1943: DELI⟩

Figure 5.2: The *jeep* entry in Rando (1987).

form *gippone* (vehicle larger than a jeep) appears to be current in present-day Italian, whereas the compound *jeep-girl* sounds obsolete. The entry also contains historical information about the adoption of this word during the American occupation in the post-WW2 years. This type of information is provided only when the author considers it relevant to the lexical profiling of the headword.

– *Dictionary of European Anglicisms* (2001)
Görlach's *Dictionary of European Anglicisms* (DEA) contains 3,800 Anglicisms present in 16 European languages, of which 1,600 are recorded for Italian. The aim of this project was to provide an analytical picture of the impact of the English language in Europe, on the basis of 16 European languages representing many language families: four Germanic (Icelandic, Norwegian, Dutch and German), four Slavic (Polish, Russian, Croatian and Bulgarian), four Romance (French, Italian, Spanish and Romanian) as well as four other languages (Finnish, Hungarian, Albanian and Greek). The time frame encompasses the years 1945–1995, the period of maximum penetration of the English language on a global scale. As for the criteria of inclusion, Görlach preferred to limit his word list by focusing exclusively on words that are "recognizably English", i.e., have an English form, discarding the large number of internationalisms, unless they retain strong evidence of English provenance (a question discussed in Section 3.1). The DEA is an ambitious contribution to comparative loanword lexicography. The method adopted for collecting and assembling the material for each language is meticulously illustrated in Görlach (2003) and the status of English in the selected languages is described in two separate volumes (Görlach 2002a, 2002b), setting the ground for systematic comparison.

Figure 5.3 shows the structure of the entry for *jeep*: the etymon is followed by a part-of-speech label (n. noun) and a definition taken from the *Concise Oxford Dictionary* (if there is more than one entry with the same lemma, a number indicates which one is considered). The definition is followed by a summary paragraph about the history of the lemma (this paragraph is present for some but not all lemmas): in this case, it is explained that *jeep* derives from the abbreviation G.P. meaning 'general purpose' (vehicle). What is interesting is the fact that the word *jeep* has not generated any replacement in the languages considered. What follows is a concise language-by-language presentation of the status of the lemma, in this case the status of *jeep,* including pronunciation, gender, plural formation, date of adoption and degree of acceptance, indicated by symbols and numbers. For instance, number (2) indicates that "the word is fully accepted and found in many styles and registers, but is still marked as English in its spelling, pronunciation, or morphology", whereas number (3) indicates that "the word is not (or no longer) recognized as English;

English origin can only be established etymologically".[78] The acceptance of *jeep* ranges between (2) and (3) in the languages represented. As can be noted, the information provided for *jeep* in Italian is that the pronunciation is close to the English one, gender is feminine, there is no plural (invariable plural), the decade of adoption is the 1940s, and the word is fully accepted.

jeep   *n*. I 'a small sturdy esp. military motor vehicle
with four-wheel drive'
   The car and its name (possibly from *G.P.* = *general
purpose* [vehicle]) became proverbially popular after
1945 as ↑*jeans* did. The word is used universally and
has not sparked off any native replacements.
Ge [dʒ-/ʃ-] M, pl. -*s*, 1940s (2)   **Du** [=E] C, 1940s (2)
Nw [jiːp/jep/=E] M, pl. -*er*, 1940s (2)   **Ic** *jeppi* [jɛhpɪ]
M, pl. -*ar*, 1940s (3)   **Fr** [(d)ʒip] F, 1940s (2)   **Sp**
[=E/jip] M, pl. -*s*, 1940s (2)   **It** [=E] F, pl. Ø, 1940s
(3)   **Rm** <=E>/*jep* [=E/ʒep] N, 1960s, via Fr (2/3)
Rs *dzhip* M, pl. -*y*, mid20c (3)   **Po** [dʒip] M, mid20c
(2)   **Cr** *džip* M, pl. -*ovi*, mid20c (2)   **Bg** *dzhip* M, pl.
-*al-ove*, mid20c (2)   → -*ka* F   **Fi** *jeeppi* mid20c
(3)   **Hu** *dzsip* [dʒip] pl. -*ek*, mid20c (3)   **Al** *xhip* M,
pl. -*e*, mid20c (3)   **Gr** *tzip* N, mid20c (2) → *tzipaki* N

Figure 5.3: The *jeep* entry in the DEA (2001).

– *Parole straniere nella lingua italiana* (2003)
Parole straniere nella lingua italiana by De Mauro and Mancini (2003; first edition 2001) is an autonomous dictionary of foreign words, based on the unabridged *Grande dizionario italiano dell'uso* or GDU (De Mauro 1999/2007). The macrostructure of this dictionary contains all the words that are marked as 'exoticism' [ES] in the GDU, which amount to 11,104 foreign words. For this reason, the number of Anglicisms recorded in this dictionary is very high, drawing on a wider time frame and an unabridged word list, encompassing all the layers of the lexicon (see footnote 75).

---

78 The classification scheme for the 'degree of acceptance' is as follows: "- the word is not known but a calque or another native equivalent is provided [. . .]; ° the word is known mainly to bilinguals and is felt to be English [. . .]; ø the word is known but is a foreignism – that is, it is used only with reference to British or American contexts [. . .]; 1 the word is in restricted use in the language [. . .]; 2 the word is fully accepted and found in many styles and registers, but is still marked as English in its spelling, pronunciation, or morphology; 3 the word is not (or no longer) recognized as English; English origin can only be established etymologically; 4 the word is identical or nearly identical to an indigenous item in the receptor language, so that the borrowing takes the form of a semantic loan only [. . .]; 5 the word, as far as the individual language is concerned, comes from a source other than English." (Görlach 2001: xxiv).

As for numbers, non-adapted Anglicisms are 5,510. Some important specifications are applied to the structure of this dictionary. The main A–Z word list contains only non-adapted borrowings, whose orthography or pronunciation make the words recognizably foreign. Two appendices provide lists of Italianized derivatives and words historically connected to foreign languages, and phraseologisms. Appendix I has words mainly of neo-Latin origin such as *anglicismo* (!), *abolizionismo*, c*oalizione*, *conformista*, *deterrente*, *editoriale*, *eurovisione*, *impatto*, *manageriale*, *microfono*, *omosessuale*, *placebo*, *preistorico*, *realista*, *romantico*, *telefono*, etc. but also adapted Anglicisms from a non-Latin English root such as *tabloide*, *quacchero* and *behaviorismo*. Appendix II contains phraseologisms (*polirematiche*) built on foreign models such as *caccia alle streghe* (witch-hunt), *cane lupo* (wolf dog), *lavaggio del cervello* (brain washing), etc. In sum, De Mauro draws a clear-cut distinction between direct borrowings (the 'real' foreign words, according to the author), on the one hand, and Italianized adaptations, derivations, and phraseological calques, on the other.

    Most 'foreign words' recorded in this dictionary belong to the technical-specialized domains and are marked as [TS] (see Table 5.2). However, not all foreign words belong to this share of the lexicon. The words *bar*, *film* and *sport* are labelled as 'fundamental' [FO], that is, they belong to the core 2,000 words of Italian. The entry *jeep* (Figure 5.4) shows that this word, despite its foreign look, is classified as common [CO], the share of vocabulary known by speakers with a medium-high level of education. The headword is followed by the pronunciation in IPA phonetic transcription, grammatical class (s.f.inv., i.e., substantive, feminine, invariable), the layer of use (common), the date of adoption, the language of origin (ingl.), form and pronunciation. Before the definition, a further specification is given of its origin (from the acronym G.P. meaning 'general purpose' ['vehicle']). Note that the reference to the military field is not openly made explicit here, although the term 'camionetta' refers to vehicles used by the army or by the police, and a generic extension to any off-road vehicle is given. The derivative *jeeppone*, made up of the English base and an Italian suffix (-*one*, expressing large dimension), is given. However, a check in newspaper archives and corpora confirms that the Italianized form *gippone* is more frequent than *jeeppone*. The variant *gip* is still current, but rare.

– Zingarelli 2022

Zingarelli 2022 is a general dictionary of Italian, whose first edition was compiled by Nicola Zingarelli in 1917. It has recently celebrated its 100th anniversary, and 12 editions have been issued over time. The publisher Zanichelli is one of the oldest in Italy, specializing especially in textbooks for school, university and professional books, and is considered the top publisher of monolingual, bilingual and specialized dictionaries. Their sources are diversified and handled by a huge editorial team. Since 1994 *Zingarelli* dictionary has been updated every year and is well-known for

jeep /dʒip/ s.f.inv. ⚆ [1943; ingl. *jeep* /dʒi:p/, dalla pronuncia delle lettere *G. P.*, iniziali della loc. *general purpose* "per tutti gli usi", sott. *vehicle* "veicolo", per indicare la versatilità del veicolo] grossa camionetta scoperta a quattro ruote motrici, predisposta per la marcia su terreni impervi | estens., qualunque vettura fuoristrada **Der.** jeeppone **Var.** gip.

Figure 5.4: The *jeep* entry in De Mauro and Mancini (2003).

its liberal policy of inclusion of neologisms and foreignisms, in constant competition with other best-selling dictionaries on the market. It also contains literary, archaic and technical vocabulary to help with the comprehension of a wide range of academic, professional and specialized texts. The 2022 edition contains 145,000 entries.

As for Anglicisms, *Zingarelli 2022* allows the extraction of Anglicisms by selecting the option 'Lingua' ('language') and then the English form. In this way all entries marked as 'vc. ingl.' (i.e., English lemma) or 'loc. ingl.' (i.e., English phrase) are displayed. This search yielded 3,990 lemmas. More inclusive searches are also possible: by selecting the abbreviation 'ingl.' in the etymology section, also words related to English such as calques and eponyms are retrieved. This search yields 4,150 lemmas.[79]

Figure 5.5 shows the entry for *jeep*. The first thing to notice is that Jeep is spelled with a capital letter and recorded as a registered trademark,[80] followed by the English phonetic transcription and a link to an audio file. In square brackets the etymological information is indicated, in this case the origin of the word from the initials G. P. and the year of adoption. The grammatical label (s.f.inv., i.e., substantive, feminine, invariable) is followed by the definition, which indicates that this word was originally used in the military field. A cross reference is given to the orthographic variant *gip* (meaning 1), which has a separate entry (including *gippone*).

– *Nuovo Devoto*-Oli 2022
Nuovo Devoto-Oli 2022 is one of the most definitive general dictionaries of Italian, printed for the first time in 1971. It has had several new editions and revisions

---

79 The increase in the number of English borrowings in *Zingarelli* dictionary (vc. ingl., i.e. English lemma) has been steady, as is shown by the following figures: 2000 edition (2,055), 2004 edition (2,219), 2006 edition (2,318), 2016 edition (2,761), 2020 edition (2,927). The etymology section yielded 4,080 entries in the 2021 edition and has risen to 4,150 in the 2022 edition (the current one at the time of writing).
80 Jeep® was registered as a trademark in 1987 by the *Chrysler Group* and is now part of the Dutch multinational company *Stellantis*.

FLESSIONE **SILLABAZIONE:** Jeep

Jeep® / *ingl.* ◀ dʒiːp /

[lettura della sigla G.P., da *g(eneral) p(urpose) (car)* '(veicolo di) uso generale' ☼ 1943]

s. f. inv.

- autovettura scoperta, potente e molto robusta, adatta ai terreni difficili, originariamente di dotazione militare **CFR.** gip (1)

Figure 5.5: The *jeep* entry in *Zingarelli 2022.*

under the direction of authoritative linguists, and has been updated yearly since 2006. It is available in paper and digital formats, also for access through mobile devices. This dictionary covers a large variety of registers to allow comprehension of literary, professional, scientific and technical texts. The number of lemmas in the 2022 edition is 110,000. This dictionary is open to neologisms but, at the same time, is vigilant about stylistic correctness. A specific section, named *"Per dirlo in italiano"* ["To say it in Italian"] examines some recorded Anglicisms and their borrowing history, and suggests "Italian equivalents for unnecessary English words, difficult to understand or pronounce" ["alternative italiane alle parole inglesi superflue, difficili da capire o pronunciare"]. The advanced search option allows for the retrieval of lemmas according to category (field), date, etymology, grammatical class and register. Setting the search to 'language' and then to 'English', the database yields 4,140 lemmas, including direct and indirect Anglicisms, the latter consisting of derivatives and scientific vocabulary of international nature.

Figure 5.6 shows the entry for *jeep*. The lemma is printed in lower case lettering, which means that it is considered a common noun. The English pronunciation is given first, followed by both the Italian one in IPA phonetic transcription and both accompanied by an audio file. Grammatical information is 's.f. invar.' (i.e., substantive, feminine, invariable). Additional grammatical information is available after extending the description window (selecting the + symbol). In this case the original English plural is provided. In the etymon section, the lemma is defined as an English word ('voce ingl.') deriving from the abbreviation G.P., which means 'general purpose', a gloss that is also translated into Italian ('per tutti gli usi'). The date of adoption in Italian is 1943.

The analysis of dictionaries carried out so far has shown that the number of Anglicisms recorded in different lexicographic works varies depending on the coverage and the target users that each dictionary has in mind. An unabridged dictionary like the GDU, on which De Mauro and Mancini (2003) was built, is meant to be a comprehensive, national dictionary of modern Italian, whereas *Zingarelli 2022* and *Nuovo Devoto-Oli 2022* are typical 'medium-sized', general purpose dictionaries,

Figure 5.6: The *jeep* entry in *Nuovo Devoto-Oli 2022.*

covering the lexicon that college students and educated users may need, and that justifies the inclusion of some technical terminology, proper names and abbreviations. Dictionaries of Anglicisms are focused exclusively on English loanwords, but the macrostructure may be limited to only direct Anglicisms, as in the DEA, or open to both direct and indirect borrowings, as in Rando (1987). Consequently, the number of Anglicisms recorded in different lexicographic works can only be relatively indicative, but not definitive.

5.2 A comparison between letter J entries

The method adopted for the selection of Anglicisms to be included in GLAD consisted in an initial step, in which candidate Anglicisms were collected from selected lexicographic sources (see Figure 3.2). Subsequently, these candidate Anglicisms were filtered through the criteria of inclusion and exclusion set by GLAD's compilers, described in Section 3.3.1 and which are briefly reported here as a reminder (Table 5.1).

The entries for letter J in the selected reference dictionaries are displayed in Table 5.2. The final column shows the entries selected for GLAD. The choice fell on the letter J because of its brevity and also to avoid the presence of Anglo-

Table 5.1: Criteria of inclusion and exclusion for GLAD.

INCLUSION		EXCLUSION	
–	Non-adapted borrowings, e.g. *jackpot*	–	Proper names and brand names, e.g. iPad
–	Adapted borrowings, e.g. *brokeraggio* (English *brokerage*)	–	Proper-name-based adjectives, e.g. Italian *shakespeariano*
–	Proper names turned generic nouns, e.g. *kleenex*	–	Frequency-boosted domestic words e.g. Italian *assolutamente* (by analogy with English *absolutely*)
–	Semantic loans, e.g. *stella* (English *star*, 'celebrity')	–	Archaisms: items obsolete before c. 1900, e.g. *cakewalk*
–	Loan translations, e.g. *tempo pieno* (English *full time*)	–	Exoticisms: lexical items from a non-anglophone speech community mediated via English, e.g. *sushi* (from Japanese)
–	Hybrids, e.g. *clown terapia* (English *clown therapy*)	–	Specialist terms not used in the general language, e.g. *proxy*
–	Pseudo-Anglicisms, e.g. *mister* (English *coach*)	–	Internationalisms based on Latin or Greek elements, e.g. Italian *telefono*.

Latinisms, which would add to the complexity of the analysis. The lemmas are accompanied by field labels following the conventions adopted by the individual dictionaries. Further semantic specification is indicated by using numbers of labels in round brackets. Round brackets are also used to indicate that the lemma

Table 5.2: Comparison of J entries in Rando (1987), the DEA (2001), De Mauro and Mancini (2003), *Zingarelli 2022*, *Nuovo Devoto-Oli 2022* and GLAD.

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
	jab (1 tech)	jab [TS]	jab (sport)	jab (sport)	jab
		jabber [TS]			
jack[1]	jack (1 tech) (playing card)	jack[1] [TS] (playing card)	jack[1]	jack[1] (giochi)	jack[1] (playing card)
jack[2]	jack (1 tech) (elettr.)	jack[2] [TS] (elettr.)	jack[2]	jack[2] (elettr.)	jack[2] (electr.)
jack[3]		jack[3] [TS] (mar.)	jack[3]	jack[3] (mar.)	

Table 5.2 (continued)

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
		jack[4] [TS] (bot.)			
		jackass [TS] (ornit.)			
		jack dempsey [TS] (ittiol.)			
		jacket[1] [TS] (abbigl.)			
		jacket[2] [TS] (tecn.)			
		jack fruit [TS] (bot.)			
	jackpot (1)	jackpot [TS] (giochi)	jackpot	jackpot (giochi)	jackpot
		jack rabbit [TS] (zool.)			
		jack russell [TS] (cinof.)			jack russell
		JAL [TS] (aer.)			
			Jacuzzi®		
jam (block)					
		jam [TS] (mus.)			
		jamboree [CO]	jamboree (gerg.)		
		jambul [TS] (bot.)			
jamming	jamming (1 tech)	jamming [TS] (telecom.)	jamming	jamming (telecom.)	
jam-session	jam session (1 tech)	jam session [TS] (mus.)	jam session	jam session (mus.)	jam session
		jangar [TS] (mar.)			
		jansky [TS] (astrofis.)			
			jay (letter j)		
			Java (inform.)	java (inform.)	

Table 5.2 (continued)

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
jazz	jazz (3)	jazz [CO]	jazz	jazz (mus.)	jazz
JATO					
	jazz-band (2)	jazz-band [CO]	jazz-band	jazz-band (mus.)	jazz band
(jazzista)	(jazzista)		(jazzista)	(jazzista)	jazzista
(jazzistico)	(jazzistico)		(jazzistico)	(jazzistico)	(jazzistico)
(jazzman)	jazzman (1)				jazzman
		jazz-rap [TS] (mus.)			
		jazz-rock [TS] (mus.)		jazz-rock (mus.)	jazz-rock
		jazzy [TS] (mus.)			
jean					
jeans	jeans[1] (trousers) (3)				
		jeans[1] [CO] (fabric)	jeans[A] (fabric)	jeans (tess.)	jeans (noun) (fabric)
			jeans[B] (agg.) (fabric)		jeans (agg.) (fabric)
	jeans[2] (denims) (3)	jeans[2] [CO] (clothing)	jeans[C] (pl.) (abbigl.)	jeans (abbigl.)	jeans (noun) (denims)
	(jeanseria)				
jeep	jeep (3)	jeep [CO]	Jeep®	jeep	jeep
		jenny [TS] (tess.)		jenny (tess.)	
		jequirity [TS] (bot.)		jequirity (bot.)	
jersey[1]	jersey[1a] (pullover) (2)				
jersey[2]	jersey[1b] (fabric) (2)	jersey [TS] (tess.)	jersey	jersey (tess.)	jersey
Jesus revolution					

Table 5.2 (continued)

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
jet	jet (aircraft) (2)	jet [TS] (aer.)	jet	jet (aer.)	jet
		jet executive [TS] (aer.)			
	jet lag (1)	jet lag [TS] (aer.)	jetlag	jet lag	jet lag
	jet liner (1)	jet liner [CO]	jet liner		
(jet-set)	jet set (1 mod)	jet set [CO]	jet set	jet set	jet set
(jet society)	jet-society (1)	jet society [CO]	jet society	jet society	jet society
		jet stream [TS] (geofis.)		jet stream	
		jewel box [TS] (inform.)	jewel box		
jigger[1]	jigger (1 tech)	jigger[1] [TS] (tess.)	jigger[1] (appliance)		
jigger[2]		jigger[2] [TS] (tecn.)			
		jigger[3]	jigger[2] (measurement unit)		
		jigsaw [TS] (giochi)			
jitterbug					
	jingle (1 rare)	jingle [CO]	jingle	jingle (pubbl.)	jingle
		jitter [TS] (elettr.)			
			jive (mus.)		
	job (1 tech)	job[1] [CO]	job[1]	job[1] (econ.)	
job (inform.)		job[2] [TS] (inform.)	job[2] (inform.)	job[2] (inform.)	
		jobber[1] [TS] (fin.)	jobber[1]	jobber (fin.)	
		jobber[2] [TS] (comm.)	jobber[2]		

Table 5.2 (continued)

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
				job center	job center
		job creation [TS] (econ.)			
		job description [TS] (ammin. az.)			
		job evaluation [TS] (amm. az.)			
		job on call [TS] (dir. lav.)		job on call (econ., dir.)	job on call
					job placement
				job rotation (amm. az.)	
	job sharing (1 tech)	job sharing [TS] (amm. az.)		job sharing (econ., dir.)	job sharing
				jobs act (polit.)	jobs act
	joker (playing-card)				
jockey	jockey (1 tech)	jockey [TS] (sport)	jockey[1] (horse riding)	jockey[1] (sport)	
			jockey[2] (in cards)	jockey[2] (giochi)	
			jockey[3] (disc-jockey)		
	jockey cap (1 tech)				
	jodhpurs (1 tech, rare, obs)	jodhpurs [TS] (abbigl.)	jodhpurs	jodhpurs	
	jogger (1)	jogger [TS] (sport)	jogger	jogger	jogger

Table 5.2 (continued)

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
jogging	jogging (2)	jogging [TS] (sport)	jogging	jogging	jogging
John Bull					
Johnny					
		join[1] [TS] (inform.)			
		join[2] [TS] (inform.)			
	joint (1 sla)	joint [CO]		joint (gerg.)	
		joint-stock [TS] (econ.)			
joint venture	joint venture (1 tech)	joint-venture [TS] (econ.)	joint venture (econ.)	joint venture (econ.)	joint venture
jolly joker	jolly (3)	jolly[1a] CO (in cards)	jolly[A1] (jolly joker)	jolly[1a] (giochi)	jolly (jolly joker)
	jolly (3)	jolly[2a] [CO] (person)	jolly[A2] (person)	jolly[3a] (person)	jolly (person)
		jolly[2b] (in sport)		jolly[2a] (sport)	
			jolly[B] (agg.)	jolly[1b] (agg.) (person)	jolly (adj.)
				jolly[2b] (agg.) (inform.)	
			Jorkyball® (sport)		
joule	joule (1 tech)	joule [TS] (fis., metrol.)	joule[A] (measurement unit)		
			joule[B] (agg.)		
		joypad [TS] (inform.)	joypad	joypad (inform.)	joypad
	joystick (2)	joystick [TS] (inform.)	joystick (inform.)	joystick (inform.)	joystick
		JPY [TS] (monet.) jug [TS] (mus.)			

Table 5.2 (continued)

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
		juggernaut[1] [TS] (relig.)			
		juggernaut[2] [TS] (polit.)			
			juggling	juggling	juggling
		juju [TS] (mus.)			
juke-box	jukebox (3)	juke-box [CO]	jukebox[1]	jukebox	jukebox
			jukebox[2] (inform.)		
			jukebox[3] (inform.)		
			jumbo[A]	jumbo[1a] (aer.)	jumbo (jet)
jumbo	jumbo (agg.) (2)		jumbo[B] (agg.)	jumbo (agg.)	jumbo
			jumbo[C] (tram)		
				jumbo[2a] (fin.)	
jumbo-jet	jumbo-jet (2)	jumbo jet [CO]	jumbo (jet)	jumbo jet (aer.)	jumbo jet
		jump [TS] (mus.)	jump (mus.)		
		jumper[1] [TS] (equit.)		jumper[1] (equit.)	
		jumper[2] [TS] (sport)			
			jumper (elettr.)	jumper[2] (elettron.)	
		jumping [TS] (sport)			jumping
				jumpsuit (abbigl.)	jumpsuit
jungle	giungla/jungla (forest)				
		jungle [TS] (mus.)			jungle (mus.)

Table 5.2 (continued)

Rando (1987)	DEA (2001)	De Mauro and Mancini (2003)	*Zingarelli* 2022	*Nuovo Devoto-Oli 2022*	GLAD
		jungle style [TS] (mus.)	jungle style (mus.)	jungle style (mus.)	
junior[1]					junior (sport)
junior[2]					junior (suite)
					junior suite
		junk bond [TS] (fin.)	(junk bond)	junk bond (fin.)	
		junk food [TS] (alim.)	junk food	junk food	junk food
		junkie [CO]			
		junk mail[1] [TS] (pubbl.)			
		junk mail[2] [TS] (inform.)	junk mail	junk mail (inform.)	junk mail
		jupiter [TS] (spett.)			
		justiciar [TS] (stor.)			
		justiciary [TS] (stor.)			
		just in time [TS] (industr.)	just in time	just in time (industr.)	
		Just Pump®			
juta, jute	juta/iuta (3)				

is quoted within another entry, e.g. (jazzista), (jazzistico) and (jazzman) are included within the entry for *jazz* in Rando's dictionary.

The first noticeable difference is that the number of Anglicisms recorded by De Mauro and Mancini (2003) is much higher than in the other dictionaries. This is mainly due to the extension of the unabridged dictionary on which this word list of Anglicisms is based. The label [TS] is indicated in most of them, except for *jamboree, jazz, jazz-band, jeans, jeep, jet liner, jet set, jet society, jingle, job, joint, jolly, juke-box, jumbo jet* and *junkie*, which are labelled as [CO] 'common' ("words

known by speakers with medium-high level of education"). Despite their status of 'common' words, *jamboree*, *jet liner*, *job*, *joint* and *junkie* have been excluded after careful examination of the available data, given that the status of 'common' words was indeed questionable.[81]

Another set of problems concerns polysemy. The lemma *jack*, for example, is recorded by De Mauro and Mancini (2003) for 4 different meanings: *jack*[1] 'playing card with a man's picture on it, worth less than a queen and more than a ten', *jack*[2] 'electric connection', *jack*[3] 'flag on a ship', *jack*[4] 'jack fruit'; the first three meanings are recorded by *Zingarelli 2022* and *Nuovo Devoto-Oli 2022*, but only two meanings, *jack*[1] and *jack*[2], have been considered current enough to be included in GLAD's word list.

The visualization of entry words and of their meanings in the selected dictionaries is extremely useful for the selection of items to include in a new database of Anglicisms, for which a particular set of criteria have been decided upon (cf. Table 5.1). Accordingly, GLAD excludes brand names such as *Jacuzzi*®, *Jorkyball*®, *Just Pump*®, but not *jeep* which has acquired a generic meaning. Highly technical terms representing specialist language registers are not included either, so that terms belonging to fields like botany (*jack*[4], *jack fruit*, *jambul*), ichthyology (*jack dempsey*), zoology (*jack rabbit*), sailing (*jack*[3], *jangar*), textiles (*jenny*, *jigger*), astrophysics (*jansky*), geophysics (*jet stream*), religion (*juggernaut*) and history (*justiciar*, *justiciary*) have been left out, whereas terms of sport, music and business can be considered for inclusion. Very specific acronyms are excluded too, such as JATO (jet assisted take-off), JAL (Japan Air Lines), and JPY (Japanese yen); proper names (*Java*); units of measurement (*joule*), foreignisms like *Johnny* (British or American recruit) or terms of encyclopedic nature like *John Bull* or *Jesus Revolution. Jupiter* (theatre) and *jay* (the pronunciation of letter j) and old-standing, fully Italianized words like *giungla* (*jungle* for the meaning of 'forest') and *juta* (fiber), from non-Anglophone speech communities and mediated via English, were eliminated too.

Careful evaluation was given to words attested in all dictionaries, a status which may confirm the currency of the lemma, unless further checks in corpora and newspaper archives proved otherwise. In the case of the letter j, Anglicisms have been grouped according to the number of dictionaries in which they are recorded. The items that have been accepted are listed first, and then the rejected items follow. In section 5.4 the lexicographic information will be combined with frequency data extracted from corpora, to refine the selection.

---

81 The attribution of variety labels to dictionary entries may be done on the basis of automatic criteria, sometimes generating dubious results. For example, if a word is recorded in specialized texts, then the label [TS] is attributed, whereas the recording of a specialized term in a general text will determine the assignment of the label [CO] (personal communication with Prof. Tullio De Mauro).

ANGLICISMS INCLUDED:

– Anglicisms recorded in 5 dictionaries: jack[1], jack[2], jam session, jazz, jeep, jersey[2], jet, jet set, jet society, jogging, joint venture, jolly (jolly joker), jukebox[1], jumbo jet;

– Anglicisms recorded in 4 dictionaries: jab, jackpot, jazz-band, jazzista, jazzistico, jeans (noun, denims), jet lag, jingle, jogger, jolly (person, noun), joystick, jumbo (adj.);

– Anglicisms recorded in 3 dictionaries: jeans (noun, fabric), job sharing, joypad, junk food, junk mail[2] (inform.);

– Anglicisms recorded in 2 dictionaries: jazzman, jazz-rock, job on call, jolly (person, adj.), juggling, jumbo (jet);

– Anglicisms recorded in 1 dictionary: jack russell, jeans (agg., fabric), jeanseria, job center, jobs act, jumping, jumpsuit, jungle (mus.), junior[1] (sport), junior (suite).

– Anglicism not recorded in any dictionary: job placement, junior suite.

ANGLICISMS EXCLUDED:

– Anglicisms recorded in 5 dictionaries: jamming, jockey[1];

– Anglicisms recorded in 4 dictionaries: jack[3], jigger[1], job[1], job[2], jodhpurs, joule[A];

– Anglicisms recorded in 3 dictionaries: jet liner, jobber[1], joint, jungle style, junk bond, just in time;

– Anglicisms recorded in 2 dictionaries: jamboree, Java, jeans (trousers), jenny, jequirity, jersey[1], jet stream, jewel box, jigger[2], jigger (measurement unit), jobber[2], jockey[2], jolly (sport), jump, jumper[1], jumper (elettr.), jungle (forest), jute;

– Anglicisms recorded in 1 dictionary: jabber, jack[4], jackass, jack dempsey, jacket[1], jacket[2], jack fruit, jack rabbit, JAL, Jacuzzi®, jam (block), jam (mus.), jambul, jangar, jansky, jay, JATO, jazz-rap, jazzy, jean, Jesus revolution, jet executive, jigsaw, jitterbug, jitter, jive (mus.), job creation, job description, job evaluation, job rotation, joker (playing card), jockey[3], jockey cap, John Bull, Johnny, join[1], join[2], joint-stock, jolly (inform, adj.), Jorkyball®, joule[B] (adj.), JPY, jug, juggernaut (relig., polit.), juju, jukebox[2] (inform.), jukebox[3] (inform.), jumbo (tram), jumbo (fin.), jumper[2], junkie, junk mail[1] (pubbl.), jupiter, justiciar, justiciary, Just Pump®.

5.3 Newspaper archives

Useful resources for the study of present-day lexis are newspaper archives, which today are available online and easily searchable. Newspapers are a primary source of information and transmission of current facts, together with other

printed, audio-visual and social media channels. The daily and periodical press has been recognized by lexicologists and lexicographers as a rich lexical storehouse for the retrieval of neologisms and Anglicisms, so that newspaper language is today the object of systematic research (Pinnavaia 2005; Rogato 2008; Demata 2014; Adamo and Della Valle 2018). Since the press covers a wide range of topics, from politics and economy to sport and entertainment, newspaper language is constantly updated with new vocabulary. Neologisms and Anglicisms are exploited in journalism because of their eye-catching function, appealing to readers and attracting their interest. Newspaper language has a key function in the process of lexical renewal, especially for specialized terms that spread to the common language and foreign words that circulate internationally and penetrate the national lexicons. Another function played by newspapers is boosting the popularity of certain words in association with particular social events, turning them into real buzzwords for a certain period of time, as, for example, the term *jobs act* discussed below.

At present daily newspapers and magazines have online editions and digital archives that can be searched in order to extract articles containing specific words and phrases and offer the possibility to list extracts in order of relevance or in chronological order, and also to group results according to specific historical timelines. The newspaper archive systematically used in this study for checking Anglicisms is that of the daily newspaper *la Repubblica*, one of the most frequently read newspapers in Italy. This archive contains all the online issues of this newspaper published since 1984. Searching for words and phrases in newspaper archives poses difficulties similar to the ones that will be pointed out with reference to corpora, namely, the need to look up multi-word units in all their possible orthographic forms (two words, hyphenated or solid form) and the impossibility to disambiguate multiple meanings. In addition, English-Italian homographs cannot be automatically distinguished (e.g. mobile, corresponding to the English adjective synonymous of 'portable' and the Italian word 'piece of furniture') (see footnote 89). In addition to polysemy, some items can be proper names of people, bands, public houses (restaurants and discoteques) or titles of songs, such as the word *jack*. In these cases, searches in archives are inconclusive.

In this study, newspaper archives proved very useful to check the currency of doubtful Anglicisms. For example, the word *jab* (in boxing) was initially excluded because it was considered too technical but was subsequently reintegrated because of its use in *la Repubblica* archive (304 articles) and because of the lack of a specific Italian equivalent, which can only be rendered with a paraphrase such as 'short, direct blow'. It is also important to point out that the number of 'hits' obtained from a newspaper archive refers to the number of articles in which the seach item appears. Not being a corpus, the numerical number of 'tokens' for the seach word cannot be retrieved. It follows that the data from the archive are not

comparable to the data extracted from the corpora. Nevertheless, the more or less intense circulation of words in newspapers can be indicative of their availability in the repertoire (at least the passive one) of readers. The observation of authentic examples from the archive has contributed to the inclusion of other Anglicisms that were not found in the reference dictionaries or were recorded by only one dictionary (see Section 5.2): for example, *job placement* (117 articles since 1999); *junior suite* (111 articles since 1996); *jack russell* (128 articles since 1992); *jobs act* (6,164 articles since 2004); *jumpsuit* (54 articles since 2008); *jeanseria* (138 articles since 1986); *job center* (193 articles since 1994), *job centre* (55 articles since 1992).[82]

5.4 Corpora: The frequency dimension at work

After having explored the quantitative dimension of Anglicisms in Italian, as recorded by different types of dictionaries, we can turn to their actual usage frequency in Italian. As aptly highlighted by De Mauro (2005) the words recorded in dictionaries are only a selection of the vocabulary circulating among speakers and is potentially unlimited in all modern languages. Most importantly, the calculation of the size of a particular lexical set only based on the number of items recorded in dictionaries is not at all meaningful without taking into consideration how frequently these lexical items are used. As for English borrowings, according to De Mauro's data, the most frequent Anglicism in spoken Italian is *okay* and, apart from that, only a few Anglicisms feature among the 5,000 most frequently used words in spoken Italian (*clan*, *goal*, *tennis*, *tram* and *whisky*). This further confirms De Mauro's conclusion that the incidence of Anglicisms is statistically insignificant, a definitive statement based on a corpus of spoken Italian, called *Lessico di frequenza dell'italiano parlato* (LIP) (1993), a collection of texts dating back to 1990–92 (De Mauro et al. 1993).[83] In this section, the role of corpus evidence for linguistic research is dealt with, using two corpora of Italian to assess the frequency of the letter J Anglicisms discussed in Section 5.2.

The use of corpora in linguistic research has revolutionized the way in which data can be gathered and analyzed, allowing linguists to access billion-sized collections of authentic language and perform sophisticated queries that would be manually impossible (Tognini Bonelli 2001) or, if necessary, build one's own *ad hoc* corpora. Corpus-based and corpus-driven investigations have already been carried

---

82 The last update of figures was done in March 2022.
83 LIP is available for consultation at http://www.parlaritaliano.it/index.php/it/volip. The VoLIP Corpus contains the transcriptions and the audio file of the LIP corpus.

out for the study of Anglicisms (Pulcini 2006; Furiassi and Hofland 2007; Laviosa 2007; Andersen 2012; Marti-Solano and Ruano San Segundo 2021). In this section, the set of candidate Anglicisms beginning with letter J were searched for in two Italian corpora, i.e., CORIS and Italian Web 2020, to obtain evidence of their currency (degree of assimilation) and representativeness (usage frequency). At present, many corpora of Italian are available,[84] but the two best suited corpora for the observation of Anglicisms are CORIS and Italian Web 2020. CORIS (*Corpus di Italiano Scritto*) was compiled at the Centre for Theoretical and Applied Linguistics of Bologna University starting from 1998. CORIS claims to be a general reference corpus of present-day written Italian, consisting of 165 million running words of texts published from 1980 to 2021. It is freely available online,[85] and is updated every three years by means of an in-built monitor corpus (last update in 2021).[86] Italian Web 2020, also known as itTenTen20, is a web-based corpus, consisting of 14.5 billion words and available on the Sketch Engine platform.[87]

A pilot study on the use of these corpora for the study of Italian Anglicisms was carried out to test their validity and the solidity of the statistical data obtained (Lukasik and Pulcini 2021). The results showed that there was an unexpected match between CORIS and Italian Web 2016 (a previous version of Italian Web 2020, containing 4.9 billion words), despite the great differences in size and content, a result that is partly confirmed by the data on letter J Anglicisms listed in Table 5.3. The major difference between these two corpora is their design, with Italian Web 2016 created by automatic means on the basis of Web-based texts, and CORIS based on specific text genres. The search interface is not very user-friendly in CORIS, relying on its own CQL (corpus query language) and lacking basic tools such as the word list option. By contrast, the Sketch Engine platform, which hosts both Italian Web 2016 and 2020, offers a wide range of easy-to-use tools for word list and concordance search, collocation generation, sorting and filtering options, which are crucial for sense disambiguation of homonyms, as will be explained below. As regards differences in the relative frequency recorded in the two corpora, they were ascribed to the problem of corpus balancing. CORIS is a balanced corpus, i.e., the texts contained represent different textual genres,[88] whereas Italian Web

---

84 Available at: https://accademiadellacrusca.it/it/contenuti/banche-dati-corpora-e-archivi-tes tuali/6228.

85 Available at: http://corpora.dslo.unibo.it/TCORIS.

86 The sampled texts in CORIS represent six textual varieties, i.e. press, fiction, academic prose (the humanities, science), legal and administrative prose, ephemera (letters, leaflets, instructions) and miscellanea. These six sub-corpora can be queried separately.

87 Available at: https://www.sketchengine.eu.

88 https://corpora.ficlit.unibo.it/coris_itaProgett.html (November, 2022).

2016 contains online texts retrieved from the web with no sampling criteria. Another important observation emerging from the pilot study regards the values recorded for absolute frequency and relative frequency, that is, whether these can be considered valuable measures of 'currency' to be used for inclusion or exclusion in the process of lemma selection for a dictionary or a database like GLAD. Whereas relative frequency is indicative of a higher or lower incidence of an Anglicism in the language, it is not advisable to set an arbitrary threshold for inclusion or exclusion. In the data referred to in the present work, for example, the item *job on call*, which scores zero in CORIS and 0.03pmw (per million words) in Italian Web 2020 was included in GLAD on the basis of other linguistic and extra-linguistic criteria (see below).

The most traditional approach to the use of corpora for lexicographic purposes is the corpus-based one, i.e., starting from a list of candidate words, such as Anglicisms, and find out about their frequency and usage context. A preliminary selection of letter J candidates was done, excluding proper names, brand names, archaisms, exoticisms, specialist terms and acronyms (see Table 5.1). The data collected for letter J lemmas are reported in Table 5.3 (frequency pmw from CORIS and from Italian Web 2020), indicating absolute frequency (AF) and relative frequency (RF pmw) side by side, to allow comparisons. The candidate Anglicisms excluded from GLAD are shaded in grey. They are understandably positioned at the bottom of the lists, among the less frequent items. Before single cases of inclusion or exclusion are discussed, two types of obstacles must be mentioned as far as corpus search is concerned. First of all, candidate items having alternative spellings (including capitalization and variant orthography) must be typed in individually in the search box, so that, for example, *job center* must be looked up as two separate words (*job center*), as a solid one (*jobcenter*) or as a hyphenated compound (*job-center*), and also with the alternative spelling *job centre*, *jobcentre* and *job-centre*. At the moment, no tools are available to make alternative spelling searches automatic.

The second type of problem regards word sense disambiguation and homonymy,[89] which was already pointed out for the four different meanings of the Anglicism *jack*. If the number of hits is manageable (e.g. 57 hits for *jack* in CORIS), the desired meaning can be identified manually, so that 25 results of *jack* 'electric connection' were isolated in CORIS. If the number of hits is too high for a manual count (e.g. 202,918 hits for *jack* in Italian Web 2020), the selection of available filters may be a solution. The procedure applied consists in extracting collocations

---

89 This problem is particularly serious in case of Italian-English homographs, such as, for example, the Italian word *mobile* ['mɔbile] (a piece of furniture, among many other meanings in Italian) and mobile ['məʊbaɪl] in English. Even more demanding and time-consuming would be the identification of new meanings attached to Italian words (semantic loans).

to find possible filters. The lemma filter applied to *jack²* was the inclusion of 'audio' within the span of 15 words in the co-text for the meaning of 'electric connection' (4,303 hits) and for *jack¹* the inclusion of 'carta' (card) for the meaning of 'playing card' (804 hits). Inclusion and exclusion filters are indicated next to items that required it in Table 5.3. In some cases the use of a filter was too complex, as in the lemma *job* (which was excluded, see below) and *junior* (which was included, see below). Moreover, direct borrowings can be easily spotted thanks to their orthographic salience, whereas calques and semantic loans remain hidden in texts and are more difficult to identify.

Table 5.3: Absolute frequency (AF) and relative frequency (RF) of letter J candidate Anglicisms retrieved from CORIS (columns 1, 2 and 3) and Italian Web 2020 (columns 4, 5 and 6).

1	2	3	4	5	6
	CORIS AF	CORIS (RF pmw)		Italian Web 2020 AF	Italian Web 2020 (RF pmw)
jeans	2,009	12.17	jazz	359,621	24.78
jazz	1,796	10.88	junior (suite) junior (sport)	166,573	11.48
jeep	699	4.23	jeans	103,194	7.11
joint venture joint-venture jointventure	504 173 4	4.12	job[1]	50,776	3.5
jet	622	3.76	jeep	50,236	3.46
junior (suite) junior (sport)	506	3.06	jolly (able to do different things, playing different roles)	46,125	3.18
jogging	243	1.47	jobs act jobs-act jobsact job act job-act jobact	31,833 159 569 2,616 56 99	2.4
juke-box jukebox juke box	100 46 27	1.04	jam (jam session)	33,081	2.28

Table 5.3 (continued)

1	2	3	4	5	6
jolly (able to do different things, playing different roles)	167	1.01	joint venture joint-venture jointventure	21,409 5,573 81	1.86
joystick joy-stick joy stick	132 9 1	0.86	jackpot jack pot jack-pot	19,695 33 14	1.36
jet set jet-set jetset	88 34 4	0.76	jogging	19,261	1.33
jumbo	84	0.5	jersey (fabric) (filter: – new)	15,461	1.07
jazzista	72	0.43	joystick joy stick joy-stick	13,990 57 56	0.97
jet lag jet-lag jetlag	43 26 3	0.43	jazzistico	13,302	0.92
jersey (fabric)	57	0.34	junior suite juniorsuite junior-suite	11,356 102 82	0.79
jam session jam-session	53 3	0.33	jazzista	10,926	0.75
jackpot jack-pot	52 3	0,33	jukebox juke box juke-box	4,556 3,080 2,484	0.69
jingle	52	0.31	jumbo (filter: – jumbo jet)	8,764	0.60
jungle (music)	46	0.27	jacket[1] (clothing) (filter: – Full Metal)	8,579	0.59
jobcenter job center job-centre	36 4 1	0.24	jam session jam-session jamsession	8,053 478 97	0.59
jobs act jobsact	27 3	0.18	jet set jet-set jetset	5,027 1,714 326	0.48

Table 5.3 (continued)

1	2	3	4	5	6
jab (boxing)	27	0.16	jingle (filter: – jingle bells)	5,876	0.40
job sharing job-sharing	26 1	0.16	jet (filter: + privato; +charter)	5,316	0.36
jack² (electr.)	25	0.15	jet lag jet-lag jetlag	3,563 1,278 433	0.36
jacket¹ (clothing)	17	0.10	joypad joy-pad joy pad	5,284 16 25	0.36
junk bond junk-bond	15 2	0.10	jack² (electr.) (filter: +audio)	4,303	0.29
jazz-band jazz band	3 12	0.09	jumping (filter: base jumping; bungee jumping)	3,919	0.27
jazzistico	15	0.09	jazz band jazz-band jazzband	3,555 121 38	0.25
junk food junk-food	14 2	0.09	junk food junk-food junkfood	3,353 240 175	0.25
jam (jam session)	13	0.07	jazz-rock jazz rock jazzrock	1,946 1,558 104	0.24
joint	11	0.06	jack russell (dog breed) jackrussell jack-russell	2,430 11 1	0.16
joypad	10	0.06	job description job-description	2,406 13	0.16
jumbo jet jumbojet	8 1	0.05	job placement job-placement jobplacement	2,354 31 18	0.16
job¹	7	0.04	jockey (filter: – disc, – disk, – club)	2,091	0.14
jam (block)	5	0.03	jab (boxing)	1,865	0.13
jack¹ (cards)	5	0.03	jazzy	1,803	0.12

Table 5.3 (continued)

1	2	3	4	5	6
jack russell (dog breed)		5	0.03	jive	1,633	0.11
jazzy		5	0.03	jolly (filter: + *carta*)	1,338	0.09
jolly (card)		5	0.03	jumpsuit	1,376	0.09
				jump-suit	4	
				jump suit	4	
jumpsuit		5	0.03	jamming (interference)	1,157	0.08
junkie (drug dealer)		5	0.03	job center	435	0.08
				jobcenter	199	
				job-center	4	
				job centre	401	
				jobcentre	115	
				job-centre	10	
job creation	4		0.03	jack[1] (cards)	804	0.06
job-creation	1			(filter: +carta)		
jazzman	3		0.03	juggling	886	0.06
jazz man	2					
jazz-rock	5		0.03	junkie (drug dealer)	799	0.06
jazz rock	1					
jive		4	0.02	jogger (filter: – baby	814	0.05
				jogger)		
jobber		4	0.02	juju (music)	713	0.05
junior suite		4	0.02	job sharing	570	0.04
				job-sharing	90	
				jobsharing	6	
junk mail	1		0.02	jazzman	364	0.03
junk-mail	3			jazz man	52	
				jazz-man	26	
jamming (interference)		3	0.01	job on call	445	0.03
				job-on-call	11	
jet executive		3	0.01	job rotation	470	0.03
				job-rotation	26	
jet-society		3	0.01	jigsaw	208	0.01
				(filter: – puzzle)	2	
				jig-saw		

Table 5.3 (continued)

1	2	3	4	5	6
jitterbug		2	0.01	jobber	430	0.03
job placement		3	0.01	junk bond junk-bond junkbond	419 16 4	0.03
job rotation		3	0.01	jamboree (rally)	233	0.02
jogger		3	0.01	jewel box jewelbox jewel-box	286 74 14	0.02
juggling		3	0.01	jumbo jet jumbo-jet jumbojet	383 29 7	0.02
juju (music)		3	0.01	jungle (music) (filter: +musica)	294	0.02
jumping		3	0.01	jacket^2 (cover) (filter: +protettivo)	41	0.01
job^2 (IT)		2	0.01	jam (block) (filter: +sistema)	199	0.01
job description		2	0.01	jazz-rap jazz rap	27 51	0.01
jockey		2	0.01	jet executive jet-executive	122 1	0.01
joint stock		2	0.01	jetliner jet liner	60 4	0.01
jump (rhythm)		2	0.01	jet society jet-society	42 16	0.01
job evaluation		1	0.006	jitterbug	133	0.01
jazz-rap		1	0.006	job^2 (IT) (filter: +elaborazione)	151	0.01
jetliner		1	0.006	job creation jobcreation job-creation	229 9 8	0.01
jacket^2 (cover)		0	0	job evaluation	78	0.01
jamboree (rally)		0	0	joker (filter: +carta)	6	0.01
jewel box		0	0	jodhpurs	39	0.01

Table 5.3 (continued)

1	2	3	4	5	6
jigsaw		0	0	joint (filter: + canna)	49	0.01
job on call		0	0	joint stock joint-stock	128 55	0.01
jockey cap		0	0	jump (rhythm) (filter: + musica)	41	0.01
jodhpurs		0	0	jumper[1] (horse) (filter: + cavallo)	35	0.01
joker		0	0	jumper (athlete of base jumping) (filter: + atleta)	9	0.01
jumper[1] (horse)		0	0	jungle style jungle-style	52 2	0.01
jumper (athlete of base jumping)		0	0	junk mail junk-mail junkmail	206 14 9	0.01
jungle style		0	0	jockey cap	0	0

The data extracted from CORIS and Italian Web 2020 offer a different visualization of candidate Anglicisms, from alphabetical order to frequency rate. Only two lexical items score a value higher than 10pmw in CORIS, namely *jazz* and *jeans*, and only a handful score a frequency above 1pmw in both corpora, confirming that Anglicisms are low-frequency words in the general language. A corpus-based criterion of inclusion does not seem a viable solution. The final decision of inclusion or exclusion was based on all the data collected, i.e., the number of dictionaries in which the Anglicisms were recorded, the frequency score in the two corpora, the number of articles in the archive of the daily newspaper *la Repubblica* (last update done in March 2022) as well as a certain amount of 'native speaker intuition'. Some controversial instances are discussed below.

The following Anglicisms have been considered current in the general language, despite the fact that they were recorded in only 1 or 2 dictionaries (or unrecorded):

– *jazzman*: 'A man who plays jazz; a jazz musician'. CORIS: 0.03pmw. Italian Web 2020: 0.03pmw. Although the frequency of *jazzman* is low in corpora and moderate in the *la Repubblica* archive (*jazzman*: 192 articles; *jazz-man*: 10; *jazz man*: 10), the currency of *jazz* and its Italian derivatives justify the inclusion of *jazzman*. The

Italian word *jazzista* is far more current (*la Repubblica* archive: 3,261 articles) and likely to replace *jazzman* altogether.

– *jazz-rock*: 'type of music which combines elements of jazz and rock'. CORIS: 0.03pmw; Italian Web 2020: 0.24pmw; *la Repubblica* archive: *jazz-rock* (*jazz rock*, *jazzrock*) 4,218. The high frequency of *jazz-rock* in the newspaper archive and in Italian Web 2020 motivates its inclusion, contrary to *jazz-rap*, which was excluded.

– *job on call*: 'working with an intermittent schedule on a needed basis'. CORIS: 0; Italian Web 2020: 0.03pmw. Dated 2001, this false Anglicisms is modelled on the English term *on call work*, which in the USA and in Canada regards the availability of regular employees to be called to work when they are needed. In Italy the meaning is different, as it refers to a temporary form of employment. This new type of contract was introduced in the early 2000s by the Italian parliament to reorganize the job market. The Italian term for this contract is *lavoro a chiamata* or *lavoro intermittente*. *Job on call* is likely to have been coined on the Italian model, which in turn was influenced by the English one, but this is difficult to establish. The data from *la Repubblica* archive indicate that *lavoro a chiamata* is more frequently used in the press (291 articles since 1994) than *lavoro intermittente* (125 articles since 1993) and *job on call* (102 articles since 2000). It was decided to include this term because it is the denomination used in Italian official documents.[90]

– *juggling*: 'the act of throwing and catching several objects continuously so that most are in the air all the time, as an entertainment'. CORIS: 0.01pmw. Italian Web 2020: 0.06pmw. The inclusion of this Anglicisms is controversial. Its frequency is low in both corpora and moderate in *la Repubblica* archive. However, the art of juggling has recently attracted interest as a hobby or a skill for street artists. This neologism has not replaced the older Italian term *giocoleria*, which is far more frequent in *la Repubblica* archive (98 articles for *juggling* since 1987 vs 1,736 articles of *giocoleria* since 2000).

– *jack russell*: 'A variety or breed of small terrier'. CORIS: 0.03pmw; Italian Web 2020: 0.16pmw; *la Repubblica* archive: 128 articles. The decision to include the name of this breed of dogs is that it has become very popular in Italian families since the 1990s (first mentioned in *la Repubblica* archive in 1992).

---

90 The new types of contracts include *job on call, job sharing* and *staff leasing*. In the section "*Per dirlo in italiano*" ("To say it in Italian"), *Nuovo Devoto-Oli 2022* criticizes the use of these English terms in an official, public document of the Italian State, addressed to all citizens, stating that the use of English expressions, whose meaning is not clear to non-experts, has the euphemistic purpose of hiding the negative implications of a job that does not offer long-term perspectives for candidates.

– *job center*: 'An office where a variety of available jobs are advertized'. CORIS: 0.24pmw; Italian Web 2020: 0.08pmw; *la Repubblica* archive: *job center* 193, *job centre* 55; *centro per l'impiego* 747. The economic crisis, unemployment and social security measures taken by the Italian state are topical issues in the daily press. For this reason, quite a few Anglicisms containing the word 'job' are included in GLAD (*job on call*, *job placement*, *job sharing* and *jobs act*). Conversely, the word *job* on its own was excluded (see below). *Job center* is frequently found in the newspaper archive, but is likely to give way to the Italian word *centro per l'impiego*, which is far more frequent.

– *jobs act*: 'name of a reform in the job market'. CORIS: 0.18pmw; Italian Web 2020: 2.4pmw. The inclusion of this term is controversial because of its status of a proper noun denoting a specific Italian reform,[91] although it is often written in small letters and partly used in a generic sense. *Jobs act* is a typical term boosted by journalism. In Italian Web 2020 it ranks among the most frequent items, after *jazz*, *junior*, *jeans*, *jeep* and *jolly*, scoring 6,164 articles in *la Repubblica* archive. Its currency is likely to decline in the future or return to a full status of proper noun.

– *jumping*: 'a sport involving the action of jumping from a certain height (such as a bridge, precipice, etc.) while secured by an elasticated rope attached to the ankles'. CORIS: 0.01pmw; Italian Web 2020: 0.27pmw. The frequency of this term is moderate, but the decision to include it is based on the fact the it is generally known by speakers as part of the compounds *bungee jumping* and *base jumping* (both terms included in GLAD), which are very popular sports in Italy. *La Repubblica* archive has 689 articles of *jumping*, as well as 225 of *bungee jumping* (from 1994) and 79 of *base jumping* (from 2000).

– *jumpsuit*: 'a one-piece garment of combined trousers and jacket or shirt'. CORIS: 0.03pmw; Italian Web 2020: 0.09pmw. The name of this garment is little current in corpora and moderately current in *la Repubblica* archive (54 articles since 2008). Anglicisms in the field of fashion are rapidly spreading on the global market, which motivates the inclusion of neologisms like *jumpsuit*, whose success deserves monitoring.

---

91 A reform of the job market was introduced in Italy in 2014 and was named *Jobs act.* A note on *jobs act* included in the section "*Per dirlo in italiano*" ["To say it in Italian"] of the dictionary *Nuovo Devoto-Oli 2022*, explains that this expression derives from a similar reform introduced in 2012 in the USA by President Barak Obama. JOBS was the acronym of *Jumpstart Our Business Start-ups.* Ignoring the real meaning of this abbreviation, and the wordplay on which it was built, in Italy this expression created some confusion, leading to its modification into *job act* or *job's act.* The equivalent Italian generic phrase 'legge sul lavoro' is proposed by the dictionary, although the impact of the Anglicism conveys a stronger impact and an aura of change and modernity.

– *junior*: (adj.) this Latin word (see 3.4.1) is used as an adjective in different contexts: in sport, 'of athletes between the age of 16 and 21, that are allowed to take part in competitions for this age group', and also in business, 'of less standing or more recent appointment'. *Junior* is the comparative form of Latin *iuvĕnis* 'young' and can be used as a proper name to indicate a younger member in case of homomymy within the same family. This practice is alien to Italian society. In sport the plural *juniores* is often used, in which case the word is a Latinism. The pronunciation of *junior* is anglicized by some speakers–[ˈdʒuːnjə] vs [ˈjunjor]–which reflects an English influence (cf. 4.1.2). *Junior* is nor recorded as an Anglicisms in any dictionary, except for Rando (1987). On the basis of its widespread English pronunciation, *junior* has been included in GLAD in the sports context and as a short form of *junior suite*. The occurrences of *junior* in the corpora and in the newspaper archive are very high and specific senses are difficult to isolate.

– *job placement*: 'a service for finding a suitable job for someone, especially a temporary job for a student or unemployed person'. CORIS: 0.01pmw: Italian Web 2020: 0.16pmw. This compound with *job* has become very common in Italian universities. It is quite frequent in *la Repubblica* archive (117 articles since 1999). Although no dictionary has included this word yet, its recent spread, especially in Italian universities for helping graduates to enter the job market, motivates its inclusion in GLAD.

The following words, though recorded in 5 and 4 dictionaries, have been excluded from GLAD after careful examination of the data available:

– *jamming*: 'interference in a radio or electronic signal which prevents it from being received or heard clearly'. CORIS: 0.01pmw; Italian Web 2020: 0.08pmw; *la Repubblica* archive: 97 articles (but overlap with other meanings). This technical meaning is overridden by other meanings, especially by 'improvising jazz music', but none appears to be frequent enough to be included.

– *jockey*: 'a professional rider in horse-races'. CORIS: 0; Italian Web 2020: 0.14pmw. In horse riding, the Italian word *fantino* is much more frequent compared to the Anglicism, which has been excluded.

– *jigger*: this term is highly polysemic in English; in Italian it is recorded for three technical meanings, namely 'appliance to die textiles', 'high frequency transformer' and 'unit of measurement'. It is difficult to pin down separate meanings in corpora and archives. For this reason it has been excluded from GLAD.

– *job*: this is a highly polysemic word in English, meaning 'a paid position for employment'. This word is hardly ever used in Italian on its own. It occurs as an

element of various compounds such as *job description*, *job placement*, *job sharing*, *job on call*, etc. This item's combination with other lexical items makes it difficult to isolate in corpora and archives. For this reason it has been excluded from GLAD.

– *jodhpurs*: 'long breeches for riding'. CORIS: 0; Italian Web 2020: 0.01pmw; *la Repubblica* archive: 0 articles. This word appears to be obsolescent.

5.5 Roundup

In this chapter, the aim was to explore the notion of the 'Anglicization' of Italian as an open question rather than a given fact. To measure the influence of English, the observation started from the quantitative evidence provided by dictionaries. The number of Anglicisms recorded in dictionaries varies greatly, depending on size, scope and criteria of inclusion, ranging from approximately 1,600 non-adapted Anglicisms in the DEA to 5,510 non-adapted forms in De Mauro and Mancini (2003). General, medium-sized dictionaries of Italian such as *Zingarelli 2022* and *Nuovo Devoto-Oli 2022* contain about 4,000 lemmas of English origin, including non-adapted and adapted lexical items. A comparative illustration of letter J Anglicisms was carried out, on the basis of which candidate Anglicisms for GLAD were selected. The next step was to look up candidate Anglicisms in a newspaper archive (*la Repubblica*) and in two corpora of Italian, CORIS and Italian Web 2020, to verify their currency, frequency, orthographic forms and meanings, a task that is time-consuming owing to the variable orthographic forms of Anglicisms, which must be individually checked, and the problem of sense disambiguation, which must be done manually. The corpus-based query proved that Anglicisms are low-frequency items in Italian, as only a few Anglicisms of the letter J candidates score a relative frequency that is above 1pmw. Data-based figures suggest that relative frequency alone is not enough to set a benchmark for inclusion or exclusion. What emerges from the research carried out so far is that it would be necessary to compile specialized corpora in order to collect more solid data on Anglicisms in specialized domains. Furthermore, user-oriented register variation in the use of Anglicisms appears to have been totally unexplored so far. If the majority of Anglicisms belong to specialized domains, also their use is confined to certain categories of speakers, such as journalists, economists, scientists and IT experts, when they are engaged in professional communication. New data on the sociolinguistic distribution of Anglicisms in spoken and written registers may indeed contribute to place the supposed 'invasion' of Anglicisms in a more objective perspective.