Insulting Nationality Words in Some British and American Dictionaries and in the BNC

Suvi Nissinen
University of Tampere
School of Language, Translation and Literary Studies
English Philology
MA Thesis
February 2015

Tampereen Yliopisto
Englantilainen filologia
Kieli-, käännös- ja kirjallisuustieteiden yksikkö

NISSINEN, SUVI: Insulting Nationality Words in Some British and American Dictionaries and in the BNC

Pro gradu –tutkielma, 109 sivua + 2 liitettä
Helmikuu 2015

Tutkielmassa tarkastellaan loukkaavien kansallisuussanojen merkintää Yhdysvalloissa ja Britanniassa julkaistuissa englanninkielisissä sanakirjoissa. Tutkielmassa selvitetään myös, kuinka korpusesimerkit tukevat sanakirjoista löydettyä tietoa.

Tutkielman teoriaosassa tarkastellaan sanojen merkityksiä ja erityisesti loukkaavuutta, loukkaavuuden merkintää sanakirjoissa, loukkaavia kansallisuussanoja sekä tutkimukseen sisältyvien kansallisuuksien keskinäisiä suhteita. Teoriaosassa todetaan muun muassa, että loukkaavuus on vahvasti kontekstisidonnaista. Lisäksi todetaan, että loukkaavuutta voidaan sanakirjoissa merkitä usein eri tavoin, ja että kansallisuussanojen mahdollista loukkaavuutta painotetaan sanakirjoissa nykyään enemmän kuin ennen.

Tutkimusmateriaalina käytetään kahtakymmentä sanakirjaa, 37 kansallisuussanaa (jotka ovat britteihin, yhdysvaltalaisiin, irlantilaisiin, ranskalaisiin, saksalaisiin ja meksikolaisiin viittaavia sanoja) ja yhtä korpusta. Sanakirjoista kymmenen on yleissanakirjoja, kahdeksan opiskelijoille tarkoitettuja, ja kaksi slangisanakirjoja. Kansallisuudet jaoteltiin tutkimusta varten kolmeen eri ryhmään.

Analyysiosassa selvitetään, kuinka sanakirjat merkitsevät potentiaalisesti loukkaavia kansallisuussanoja; kuinka usein ja millä tavoin sanakirjat merkitsevät näitä sanoja loukkaaviksi, kuinka eri sanakirjatyypit ja yksittäiset sanakirjat eroavat toisistaan tässä suhteessa, ja millaisia eroja eri kansallisuuksien loukkaavuuden merkitsemisessä on. Korpusanalyysissä tutkitaan, ovatko sanakirjoissa yleisesti tai harvoin listatut sanat yleisiä ja harvinaisia myös korpuksessa, ja kolmen sanan loukkaavuutta korpusesimerkeissä tutkitaan myös tarkemmin.

Tutkimuksessa todettiin, että jonkinlainen varoitus loukkaavuudesta oli sanakirjoissa mukana 75 % sanoista. Sanakirjatyyppien ja yksittäisten sanakirjojen välillä oli eroja; useimmiten loukkaavuus mainittiin brittiläisissä vieraskielisille opiskelijoille tarkoitetuissa sanakirjoissa, ja harvimmin slangisanakirjoissa. Useimmiten loukkaavuus ilmaistiin kielenkäyttömerkitsinten (engl. *usage labels*) avulla, mutta myös sanojen määritelmiä, kielenkäyttökommentteja (engl. *usage notes*) ja eri keinojen yhdistelmiä käytettiin apuna loukkaavuuden merkitsemiseen. Kansallisuuksien välillä oli myös eroja; useimmiten loukkaaviksi merkittiin meksikolaisiin viittaavat sanat, ja harvimmin yhdysvaltalaisia merkitsevät sanat. Korpusinformaatio vastasi melko hyvin sanakirjoista saatua tietoa, mutta erojakin löytyi. Tutkimus osoitti, että nykyistä yhdenmukaisempi ja paikoin myös yksinkertaisempi merkintätapa loukkaaville sanoille olisi monesti tarpeen sanakirjoissa.

Avainsanat: loukkaavuus, kansallisuussana, sanakirja, korpus

Table of contents

1. Introduction **1**

2. Dimensions of word meaning **5**

2.1 Descriptive and non-descriptive meaning 5

2.2 Denotation, reference and connotation 10

2.3 Insults and offensive words 12

2.4 Ethnic slurs and insulting nationality words 16

3. Usage in dictionaries **19**

3.1 Indicating meaning and usage in dictionaries 19

3.2 Marking insult in dictionaries 22

3.3 Ethnophaulisms in dictionaries 26

4. Historical background: the relations between the nationalities examined **29**

5. Materials and methods **33**

5.1 Dictionaries studied 33

5.2 Labels in the dictionaries studied 37

5.3 The corpus used 43

5.4 Words examined 44

5.5 Methods employed 50

6. Dictionary findings **54**

6.1 Frequencies of indications of offensiveness 54

6.1.1 Frequencies for dictionaries 55

6.1.2 Frequencies for nationalities 63

6.2 Details of the dictionary findings 68

6.2.1 Findings for Group 1 69

6.2.2 Findings for Group 2 77

6.2.3 Findings for Group 3 80

7. Corpus findings **88**

7.1 Frequencies of the nationality words in the BNC 88

7.2 *Boche*, *Pom* and *Yank* in the BNC 92

7.2.1 *Boche* in the BNC 92

7.2.2 *Pom* in the BNC 93

7.2.3 *Yank* in the BNC 94
8. Discussion **96**
9. Conclusion **100**
References **105**
A. Primary sources 105
B. Works cited 106
Appendices **110**

1. Introduction

In the era when hate speech has become criminalised, it is important to be conscious about the potential offensiveness of words relating to nationality, ethnicity and race. Xenophobia —the fear and dislike of people from other countries— is not a new phenomenon, but the public opinion on minorities has changed over time. Nowadays, discrimination is punishable by law as governments have started to recognise the equal status of minority groups. Hughes (2006, 149) argues that "the language of ethnic insult has become genuinely taboo" and thus language users have to be cautious of which words not to use. Offensive and racist words cause harm to individuals and groups and can also damage the whole society (Hudson 2003, 46).

According to Landau (2001, 232), dictionaries reflect changes in society. Therefore, dictionary editors have also started to pay special attention to insults when compiling dictionaries (ibid.). Sometimes ethnic words are omitted from dictionaries in order to avoid offending anyone (Béjoint 2010, 207). Henderson (2003, 53), for instance, mentions a protest against Merriam-Webster Inc. that took place in 1997, in which the word *nigger* was demanded to be removed from their dictionaries. However, dictionaries reflect the world views of a society and omitting words that refer to a particular group is "a way of denying their existence" (Béjoint 2010, 207). According to Béjoint (ibid.), dictionaries have started including more of potentially offensive ethnic words nowadays. This makes it even more important for lexicographers to pin down the meaning and connotations of these words and to warn the reader about their potential offensiveness.

The topic of this study is insulting nationality words in some British and American dictionaries and in the British National Corpus, with the main focus being on dictionaries. I will examine potentially insulting words that refer to different nationalities and whether and how they are marked as offensive in dictionaries. Moreover, I will pay attention to possible differences in the treatment of different nationalities. The study will also shed light on potential differences between various types of

dictionaries. In addition, these dictionary findings will be compared to corpus data to see whether actual usage is in conformity with them.

This study will detail the ways dictionaries treat potentially insulting nationality words and thus it will also shed light on potential shortcomings in dictionaries when it comes to marking insulting nationality words. For this reason, the study will be useful for lexicographers. Furthermore, it will also be of use for dictionary users: the study will examine the differences in the labelling systems between various dictionaries and thus provide dictionary users with directions on how to interpret labels and other indications of offensiveness (or the lack thereof) in dictionaries.

Insulting nationality words in dictionaries have been studied to some extent earlier (see Campbell, 2006; Henderson, 2003; Norri, 2000 and Wachal, 2002, for instance), but to my knowledge no extensive study has been carried out on insulting nationality words in monolingual English dictionaries where several nationality groups would have been compared to each other. Previous studies have mostly included a smaller selection of words or even focused on a single word, and fewer dictionaries have been investigated. This study will be a more extensive one and include several nationalities across the world. What is more, ethnic words relating to race have been examined to a great extent, but nationality words have not received an equal amount of attention from scholars. It is as important to use politically correct terms when referring to different nationalities as it is when referring to people with different skin colour.

In this study, my aim is to answer the following research questions:

1. How often and how are the nationality words examined marked as offensive in British and American dictionaries?

2. How do the labelling policies of dictionaries differ from each other?

3. Are there differences in labelling when it comes to different nationalities?

4. How does corpus data correspond to the data collected from dictionaries?

Question 1 deals with frequencies and the ways of marking potentially insulting words as offensive in dictionaries. The different ways of classifying words as insulting in dictionaries will be identified; some dictionaries use labels or usage notes while others state the insulting nature in the definition, for example. Yet other dictionaries do not always mention offensiveness at all. It will be seen how frequent these different types are. Question 2 pertains to comparing different types of dictionaries and also individual volumes. There is a great deal of variation between different dictionaries when it comes to labelling, for example (Landau 2001, 232), and this variation will be explored. It will be investigated whether certain dictionaries, such as learner's dictionaries, use the indications of offensiveness differently from general purpose dictionaries, for example. Norri (2000, 91) argues that learner's dictionaries tend to use more warnings with potentially offensive words. It will be studied whether that is the case here.

As concerns Question 3, I aim to examine whether words denoting a certain nationality are seen as more insulting than words referring to some another nationality. Rice et al. (2010, 121) argue in their study that sometimes people regard insults targeting their own group as more insulting than insults targeting some other group. However, Landau (2001, 232) states that ethnic slurs targeting groups that have been severely abused in the past, such as African Americans, are "regarded with special outrage". What is more, Norri's (2000, 83) study also suggests that the words denoting oppressed groups seem to be marked as more offensive than those denoting the oppressors of these groups (or the "non-oppressed"). Thus, it will be very interesting to see whether British dictionaries, for example, classify the words denoting the British as more offensive than the words denoting the Irish – or the other way round.

Question 4 deals with the relation between dictionary findings and corpus data. The aim is to investigate whether the inclusion and exclusion of words in the dictionaries corresponds to corpus frequencies. That is to say, whether words rarely found in the dictionaries examined are also rare in the

corpus, for example. Moreover, three nationality words will be examined in more detail. Again, it will be studied whether the corpus data corresponds to the dictionary findings. If dictionaries have indicated that a particular word is extremely offensive, for instance, it will be investigated whether the examples in the corpus support this.

The main material for the study consists of various monolingual dictionaries. The words that will be examined are words that refer to the British, Americans, the Irish, the French, Germans and Mexicans. In total, 37 words will be studied by using 20 dictionaries. The materials and methods of the study will be further discussed in Chapter 5. More British than American dictionaries will be used in the analysis, and many of the nationalities studied (such as the Irish and the French) share more history with Britain than with the United States. In addition, the corpus used in the study is a corpus of British English. Therefore, the focus of the study will be more on Britain.

This MA thesis consists of nine chapters. Chapter 2 will provide the reader with theoretical background to word meaning, insults and insulting nationality words. Chapter 3 will focus on lexicography, and more specifically on the ways of indicating usage and the potential offensiveness of words in dictionaries. Chapter 4 will introduce some historical background to the nationalities that are included in this study and also shed light on the relations between the countries that these nationalities represent. The materials and methods of the study will be discussed in Chapter 5. Dictionary findings will be presented in Chapter 6, and the corpus data will be analysed in Chapter 7. The findings will be discussed further in Chapter 8. Lastly, the conclusions of the study will be presented in Chapter 9.

2. Dimensions of word meaning

When compiling a dictionary, the essential task of a lexicographer is to write down definitions for a multitude of words and thus to capture their meaning. Word meaning consists of several components: the relation between a word and the real world, all the associations that this word has, the ways the word relates to other words and the context of the sentence it occurs in (Jackson 2002, 15). Moreover, many words are polysemous – that is to say, they have multiple meanings (ibid.). Thus, word meaning is a complex concept and it consists of numerous components.

Before discussing the issues related to lexicography and the treatment of offensive words in dictionaries, it is useful to discuss the notion of word meaning in general. Lexicographers have to be aware of the nuances between the meanings of different words, and those nuances are also important in this study. Even though two words may refer to the same nationality, for instance, only one of them may be an insulting word. Furthermore, it is essential to explain all the concepts relating to word meaning which will be used in the analysis.

In this chapter, different dimensions of word meaning will be explored. Section 2.1 will focus on descriptive and non-descriptive meaning. In Section 2.2, the concepts of denotation, reference and connotation will be discussed. Section 2.3 will shed light on insults and offensive words. In Section 2.4, the focus will be on ethnic slurs and insulting nationality words.

2.1 Descriptive and non-descriptive meaning

Word meaning can be analysed and subcategorised in various ways. As Cruse (2011, 195) notes, scholars are not unanimous when classifying word meanings into different types. According to Cruse (ibid.), there is, however, "substantial agreement" as concerns one particular type of meaning. In this thesis, Lyons' (1977) term *descriptive meaning* will be adopted in order to refer to this meaning type.

Other terms that have been used in the same sense include *referential*, *cognitive*, *propositional*, *ideational*, *designative* and *logical meaning* (Cruse 2011, 196; Lyons 1977, 51).

The theoretical framework when analysing descriptive meaning is the truth-conditional theory: descriptive meaning is linked to the truth-conditional properties of an expression (Jaszczolt 2002, 3; Cruse 2011, 195–196). That is to say, words with descriptive meaning contribute to the truth or falsity of a statement (Cruse 2011, 195–196). Descriptive meaning "can be explicitly asserted or denied and, in the most favourable instances at least, it can be objectively verified" (Lyons 1977, 50). Lyons (ibid., pp. 50–51) gives the following example of a sentence with descriptive meaning:

(1) It is raining here in Edinburgh at the moment.

As is the case with expressions carrying descriptive meaning, (1) can be asserted, denied and probably verified, too. A word such as *Edinburgh* contributes to the truth conditions: if the sentence was *It is raining here at the moment*, it would not necessarily be true that it was raining in Edinburgh.

The words examined in this thesis can be categorised into groups of words that share the same descriptive meaning; the words denoting the Irish and the words denoting Mexicans, for example. Obviously, the descriptive meaning between the words belonging to different nationality groups varies. However, the descriptive meaning may also vary between the words in the same group, such as between the words referring to Mexicans. These differences will be discussed in Section 5.4.

Even though there are differences between the meanings of words within the same nationality group, they can still be treated as synonyms, at least to some extent. Cruse (2011, 143) uses the concept *propositional synonymy* when referring to words that are interchangeable without changing the truth-conditional properties of the sentence, one word entailing the other (ibid.). Thus, propositional synonyms share the same descriptive meaning. The terminology varies to some extent: propositional synonymy has formerly been called *cognitive synonymy* by Cruse (1986, 88), and Lyons (1995, 63) uses the term *descriptive synonymy* when referring to this sense relation. Here, Lyons' (1995) term

descriptive synonymy will be adopted for the sake of consistency, as Lyons' (1997) term *descriptive meaning* is used as well.

Words that share the same descriptive meaning may differ in various ways. Cruse (2011, 143) states that the three most important differences between descriptive synonyms are related to formality, emotional state (see the discussion on expressive meaning on p. 9) and the presupposed field of discourse. Jackson (1988, 73) adds that descriptive synonyms may belong to different dialects and that they can be differentiated in terms of euphemistic value. According to Jackson (ibid.), euphemistic words are used when referring to taboo subjects "more obliquely", and thus they only cover certain semantic fields. For example, the word *intoxicated* is a euphemism, "though actually perhaps a more formal term" for drunk (Jackson, ibid.). These categories presented by Jackson (ibid.) and Cruse (2011, 143) seem to overlap; Jackson's (1988, 73) category of euphemisms, for instance, seems to overlap with Cruse's (2011, 143) category of formality. Therefore, the boundaries between some of these categories appear to be rather fuzzy.

The differences between descriptive synonyms mentioned above can be characterised as features that represent non-descriptive meaning. Cruse (2011, 196) gives the following sentence (2) as an example of both descriptive and non-descriptive meaning:

(2) Somebody's turned the bloody lights off.

In (2), the word *off* has descriptive meaning: it is crucial as far as the truth-conditions of the sentence are concerned (ibid.). The proposition would change essentially if *off* was left out or replaced with *on,* for example. However, *bloody* has no descriptive function: the 'basic' meaning of the sentence would remain the same even if the word *bloody* was deleted (ibid.). It would still be true that somebody had turned the lights off. Thus, the function of *bloody* is non-descriptive (ibid.). (2) shows that sometimes a sentence can have elements of both descriptive and non-descriptive meaning. Moreover, single words can also carry both descriptive and non-descriptive meaning (ibid., p. 201). Cruse (ibid.) gives the

word *damn* (*It was damn cold*) as an example: the descriptive meaning of *damn* is ‘extremely’, but *damn* also conveys an attitude that the neutral *extremely* does not. This second, attitudinal function of *damn*, is an example of non-descriptive meaning.

As is the case with *damn*, insulting nationality words can also be seen to carry both descriptive and non-descriptive meaning. The word pair *Kraut* – *German* can be quoted as an example. Both words refer to a person from Germany, but *Kraut* also conveys an attitude that the word *German* does not; *Kraut* is an offensive word (the *OED3*[1] s.v. *Kraut n.* and *adj.,* sense A2a). Thus, *Kraut* carries the descriptive meaning of ‘German’, but it also carries the non-descriptive meaning of offence.

Instances of non-descriptive meaning can be categorised in various ways. Nevertheless, the distinctions between these subcategories are not very clear, and for this reason many scholars have only used a single term to collectively describe all instances of meaning that have a non-descriptive[2] function (Lyons 1977, 51). For instance, terms such as *emotive*, *attitudinal*, *interpersonal* and *expressive meaning* have been used (ibid.). Lyons (1977) and Cruse (2011), on the other hand, divide non-descriptive meaning into two categories: *evoked* (Cruse) or *social* (Lyons) *meaning* and *expressive meaning*.

For Cruse (2011, 200–203), *evoked meaning* refers to variation according to dialect and register. Some words are associated with different geographical, temporal and social factors, some words occur only in particular areas of discourse, some words are only used in text messages, for instance, and some words seem more formal than others (ibid.). *Evoked meaning* refers to meaning associated with one or more such characteristics (ibid.). Lyons’ (1977) term, *social meaning*, on the other hand, is defined as “that aspect which serves to establish and maintain social relations” (Lyons 1977, 51). Cruse (2011, 203) mentions social factors as one aspect of his category of evoked meaning, but adds

---

[1] The Oxford English Dictionary, third edition; see References.

[2] Lyons (1977, 51) does not explicitly use the term *non-descriptive meaning*, but he contrasts other types of meaning with the descriptive type.

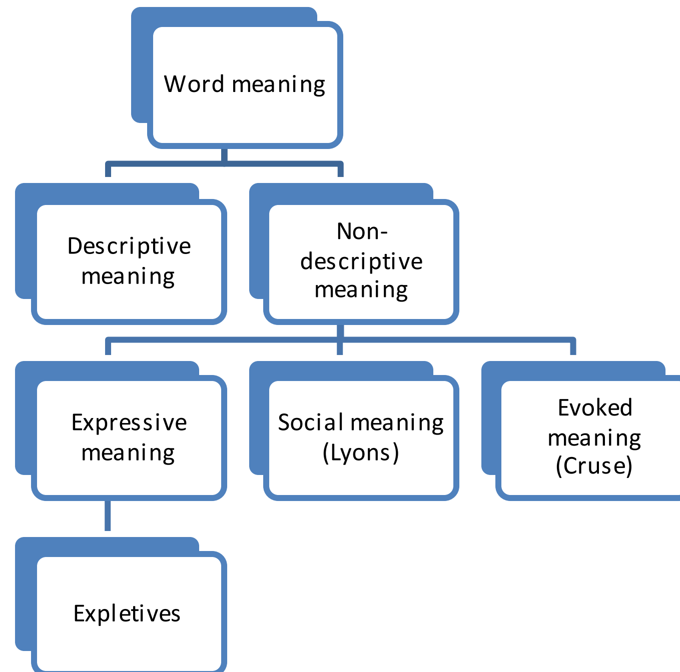
many other characteristics, such as geographical and temporal variation. Thus, Lyons' (1977, 51) concept seems to be narrower than Cruse's (2011, 203).

Both Lyons (1977) and Cruse (2011) use the term *expressive meaning* as the other category of non-descriptive meaning. Lyons (1977, 51) argues that the expressive meaning of a word is the meaning that varies according to the speaker's characteristics, but no examples of such characteristics are given. For Cruse (2011, 201), words that carry expressive meaning express an emotional state. Thus, offensive nationality words would carry expressive meaning as they depict a pejorative attitude, which can be seen as a state of emotion (feeling contempt). What is more, Cruse (ibid.) points out that the words that carry only expressive meaning and not a descriptive one can be called *expletives*. Cruse (ibid.) lists the words *wow* and *ouch* as examples.

In this case, Cruse's (2011, 201) concept of expressive meaning seems to be narrower than Lyons' (1977, 51), as only emotional states are discussed in Cruse's (2011, 201) text. Lyons (1977, 51) argues that there is overlapping in his categories of expressive and social meaning, and this seems to be the case with Cruse's (2011) categories of expressive and evoked meaning, too. It indeed seems to be the case that non-descriptive meaning is a difficult concept to subcategorise.

The following figure summarises the discussion of this section:

Figure 1. Descriptive and non-descriptive meaning



2.2 Denotation, reference and connotation

Word meaning can also be divided into *denotation*, *reference* and *connotation*. The theoretical framework which discusses these characteristics of word meaning is known as *the referential theory of meaning* (Jaszczolt 2002, 3). In the referential theory, word meaning is thought to derive from reality and the way words and longer constructions relate to situations (Saeed 2009, 24–25). *Denotation* is the meaning that refers to a whole class of things (Cruse 2011, 46). For example, the denotation of *bird* would be the whole class of birds, including all instances of birds. According to Murphy (2010, 32–33), denotative meaning is "the kind of meaning that is most directly represented in dictionary definitions of a word". Therefore, it is necessary to discuss the concept of denotation and related terms in this thesis, as the focus of the study is on dictionaries.

In addition to denotation, it is useful to discuss the term *reference*. While denotation refers to the whole class or set of things, reference "has to do with the relationship which holds between an

expression and what that expression stands for on particular occasions of its utterance" (Lyons 1977, 174). Thus, a particular 'concrete' thing that a word refers to is known as its *referent* (Cruse 2011, 46): when uttering *The dog was so kind* the referent of *dog* is one particular dog. The term *denotation*, on the other hand, can be used to refer to a class of things (ibid.): *dog* also refers to the species known as dog. As Saeed (2009, 24) notes, "referring is what speakers do, while denoting is a property of words". Terminology is again complex; as Cruse (2011, 46) points out, sometimes the word *reference* can be used to refer to either the whole class of things or to a particular entity. Jackson (1988, 50), for example, treats the terms *denotation* and *reference* interchangeably. However, in this thesis Cruse's (2011) system will be adopted.

Connotation, on the other hand, can be defined as the "semantic associations that a word has, which are not strictly part of the denotative meaning of the word" (Murphy 2010, 33). As Jackson (2002, 16) notes, "[w]hile the denotation is the straightforward, neutral relation between a word and its referent, the connotation brings in the, often emotive, associations that a word may have for a speaker or a community of speakers". For example, the words *feline* and *kitty* denote the same class of things (cat) even though their connotations differ: *feline* is a more scientific word, usually not used with domestic animals, whereas *kitty* is often seen as "something small and cute" (Murphy 2010, 33). The notion of connotation is also important when discussing potentially offensive nationality words, as it is the connotation that differentiates an insulting nationality word from a neutral one.

Allan and Burridge (2006, 29) use the term *cross-varietal synonymy* to refer to words that denote the same basic thing but which have different connotations. This concept seems to be linked to propositional synonymy, which was discussed in Section 2.1. Moreover, the concepts of non-descriptive meaning and connotation can be seen to overlap. Nevertheless, with non-descriptive meaning the focus is on the truth conditions of the whole sentence (or, how the non-descriptive meaning of a particular word makes no contribution to the truth conditions of the whole sentence). As

concerns connotation, the focus is on a particular word and its possible associations, not on the truth-conditional properties of the sentence.

Jackson (1988, 59) notes that connotations are "far more indeterminate than denotations" – they vary from one generation to another, and can be highly subjective. Even though some connotations are shared by the whole speech community, some are more restricted (ibid., pp. 58–59). Sometimes different individuals from the same community associate different connotations with a particular word (Allan and Burridge 2006, 32). Therefore, connotations depend on the context and the community or individual using the words (ibid.). For this reason, the task of a dictionary compiler is a difficult one: connotations vary from time and place to another, and thus it is difficult to pin down the potential offensive connotations of a word.

Nonetheless, as Allan and Burridge (ibid., p. 31) note, some words are unmarked in the sense that they do not carry any particular connotations. They (ibid.) give the word *dog* as an example of such case; it can be seen as a neutral word in the sense that its use is not restricted to a particular dialect or register, for instance. Furthermore, it is not an offensive word and it does not convey any other kind of attitudinal connotation, either.

2.3 Insults and offensive words

Word meanings change over time, and so does the conception of offensiveness. Social, political and historical factors have an effect on the expressions that are considered offensive (Battistella 2005, 83). Religious taboos and taboos linked to sexual obscenity have become less severe, and nowadays taboos connected with race or gender, for example, are regarded with special outrage (Burridge 2002, 201). As Burridge (2004, 57) notes, "language change typically follows social change". Henderson (2003, 54) lists occupations ending with *-man* as an example of this phenomenon. Some occupations were

reserved to men in the past, but nowadays these occupations are available to women and this can also be seen in the vocabulary (ibid.). For example, *letter carrier* can nowadays be used instead of *mailman* (ibid.).

Furthermore, in addition to changing over time, taboos vary from one culture to another. Allan and Burridge (2006, 10) give the situation of women eating with men as an example: in the late 18th century, this was regarded as a taboo among Tahitian women, but not among Captain Cook's men. As Allan and Burridge (ibid., p. 11) note, "[t]here is no such thing as an absolute taboo (one that holds for all worlds, times and contexts)".

Words can be offensive in different ways, some being truly harmful and others causing some squeamishness (Battistella 2005, 83). In addition, the same word is not equally offensive in all environments (ibid.). Avoiding coarse language can be seen as a mark of sophistication in some contexts, but, on the other hand, the use of coarse language and insults can sometimes "establish lower-order solidarity" (ibid., p. 84). Adams (2009, 200) notes that slang is offensive on many occasions, but it also has a defensive function, as "the best defense is a great offense". The notions of polite and offensive language vary according to the domain: it is different to define politeness and offensiveness in terms of press conferences or speeches than it is to define them in terms of a "common person" (Battistella 2005, 84).

The differences between in-group and out-group use of potentially offensive words should also be taken into account. A classic example would be that of the n-word. *Nigger* has traditionally been seen as an insult used by whites and targeting African Americans (Henderson 2003, 67). On the other hand, the word can also be used as a neutral or even a positive term by African Americans when referring to other African Americans (ibid.). However, Henderson (ibid.) points out that this division into two basic meanings is too narrow: *nigger* can also be used by both whites and African Americans as a counterpart for 'white trash' for someone with a darker skin. According to Henderson (ibid., p. 68),

this restricted offensive usage, referring to behaviour and socioeconomic status and not only to skin colour, is often neglected in dictionaries. Furthermore, the n-word can sometimes even be used to refer to someone who is white (Croom 2013, 190). Croom (ibid.) uses the term *non-paradigmatic derogatory use of slurs* when a slur is targeted towards someone who is not seen as the typical target of the slur.

Offensive language can be categorised in various ways. Battistella (2005, 72) divides offensive words and expressions into three groups: *epithets*, *profanity*, and *vulgarity/obscenity*. *Epithets* are slurs referring to race, ethnicity, gender, sexuality, appearance, disabilities or other characteristics (ibid.). The term *profanity* refers to religious cursing, whereas *vulgarity* and *obscenity* in their turn refer to words and expressions "which characterize sex-differentiating anatomy or sexual and excretory functions in a crude way" (ibid.). The difference between *vulgarity* and *obscenity* is "primarily a matter of degree and prurience" (ibid.), but in this study it is enough to combine these two concepts into one category. It is possible that an expression belongs to several of these three categories. Battistella (ibid.) gives the expression *God fucking dammit* as an example.

Allan and Burridge (2006, 79–85) distinguish between seven types of insults. They are listed below, with examples given by the authors (ibid.) in brackets:

1. Comparisons of people with animals that are conventionally ascribed certain behaviours (*bitch, mongrel*)
2. Epithets derived from tabooed bodily organs, bodily effluvia and sexual behaviours (*asshole, shit, fucker*)
3. Dysphemistic epithets that pick on real physical characteristics that are treated as though they are abnormalities (*baldy*)
4. Imprecations and epithets invoking mental subnormality or derangement (*retard*)
5. –IST dysphemisms (*sexist*, *racist*, *speciesist*, *classist*, *ageist* and other –IST words, such as *paki*)
6. Slurs on the target's character (*nerd*)
7. Ritual insults (all of the aforementioned types can be utilised in this competitive game. Here, insults are used to express group solidarity)

Interestingly, when listing numerous types of insults, Allan and Burridge (2006, 79) argue that "[i]nsults are normally intended to wound the addressee or bring a third party into disrepute, or both. They are therefore intrinsically dysphemistic" (see the following paragraph for the definition of *dysphemistic*). Among these insults, they (ibid.) mention skin colour. However, later they (ibid., p. 83) argue that "[r]acist terms are not intrinsically dysphemistic". With these cases they (ibid., p. 84) refer to in-group use and mention *nigger* as an example. They (ibid.) also point out that sometimes this in-group use can be extended to outsiders. An example of this is given by Kennedy (2002, 112–114), who mentions a case where a basketball coach at Central Michigan University asked his team for permission to call the players *niggers* in order to create team spirit. The team (mostly black) gave him the permission, and this way the in-group use was extended to an outsider. However, the coach was later fired because of the n-word he had used (ibid.). This shows the way insulting ethnic words are becoming a taboo in our culture.

Allan and Burridge (2006, 29) use the term *X-phemism* when discussing different ways of referring to taboo terms. X-phemisms can be further categorised into *euphemisms*, *orthophemisms* and *dysphemisms* (ibid.). A *dysphemism* is a word or a phrase that carries offensive connotations (ibid., p. 31). *Orthophemisms* and *euphemisms* can be used instead of a dysphemism if the speaker or writer wants to avoid insulting anyone (ibid., p. 32). An orthophemism is a more formal and literal choice than the more colloquial and figurative euphemism (ibid.; p. 33). Allan and Burridge (ibid., p. 32) give the words *faeces*, *poo* and *shit* as examples of an orthophemism, euphemism and dysphemism, respectively.

To conclude, it can be said that offensive language seems to be as difficult to categorise as word meaning in general. This is not surprising, as insults are one form of word meaning. The complexity of offensive language is summarised by Battistella (2005, 83):

> What seems clear overall is that the notion of offensive language is a variable one, shifting over time, relative to domain (the workplace, broadcast media, literature, political discourse, polite conversation), and affected by social, historical, political, and commercial forces. It is clear as well that the range of offensive language extends from usage that is simply offensive to the squeamish to language that is disruptive and harmful.

As Battistella (ibid.) notes, the role of context is essential, and the degree of offensiveness varies. Even though offensive language can be categorised in several ways, insults —as the language in general— are always in a state of flux.

2.4 Ethnic slurs and insulting nationality words

There are multiple ways of referring collectively to racial insults and insulting nationality words: terms such as *ethnic slurs*, *racial slurs*, *ethnic/racial epithets*, *ethnophaulisms* and *racist dysphemisms* are used. According to Croom (2013, 179), slurring terms are "used by speakers primarily to identify members that possess certain descriptive features (e.g., race) and to derogate them on that basis". Thus, *ethnic slur* derogates people based on their ethnicity and *racial slur* targets a person's race. Epithets are "various types of slurs" (Battistella 2005, 72), and thus *ethnic epithet* is another term for *ethnic slur* (and, correspondingly, *racial epithet* can be used instead of *racial slur*). *Ethnophaulisms* are "verbal ethnic slurs to refer to out-groups" (Rice et al. 2010, 118). This term was first used by Roback (1944) (Rice et al. 2010, 118). According to the *OED3*[3] s.v. *ethnophaulism n.*, the word is "[a] contemptuous expression for (a member of) a people or ethnic group; an expression containing a disparaging allusion to another people or ethnic group". In addition, the term *ethnonym* can be used to refer to ethnic words that are not offensive (the *OED3* s.v. *ethnonym n.*). Allan and Burridge (2006, 83) use the term *racist dysphemism* and mention that they "will not subclassify racists into nationalists versus ethnicists, etc.". Thus, the same word is used for slurs denoting nationality and slurs denoting skin colour or race. This policy seems to have some justification, as the issues relating to nation and race often co-occur. In this

[3] *The Oxford English Dictionary*, third edition; see References.

study, the terms *ethnophaulism* and *ethnic slur* will both be used, in order to avoid excessive repetition. They will be used to denote slurs targeting nationality and/or skin colour, even though the present study focuses on nationality words. The word *ethnic* has the meaning of "connected with or belonging to a nation, race or people that shares a cultural tradition" (*OALD*[4] s.v. *ethnic adj*., sense 1).

Xenophobia and xenophobic words are not new phenomena. As Hughes (2006, 515) argues, martial and religious rivalry can be regarded as the creators of early xenophobic attitudes. Wars have obviously been a significant source for new xenophobic words and expressions (ibid., p. 147). According to Hughes (ibid., p. 146), ethnic slurs usually develop in the following situations: migration, immigration, war, religious conflict, territorial expansion, colonialism and rivalry in business and politics.

Not all ethnic groups and nationalities are treated similarly. Verkuyten and Thijs (2010, 469) argue that different ethnic groups are not evaluated in a similar manner: hierarchies are formed and some groups are seen as more negative than others. Battistella (2005, 83) argues that "[r]elatively unassimilated groups are more likely to be targets of majority xenophobia than groups which are more assimilated into the dominant culture". Therefore, it would seem that the stranger or more alien a group of people is seen as, the more xenophobically they are treated. According to Verkuyten and Thijs (2010, 468), the distinction between in-group and out-group members maintains prejudice and discrimination. They (ibid.) note that "[p]eople tend to experience less positive affect toward members of the out-group, remember more negative information about out-group members, and are less helpful toward out-group than in-group members".

Tolerance for certain type of offensive language, such as religious and sexual expressions, has increased over time, whereas the taboo status of ethnic and racial epithets has remained or become stronger (Battistella 2005, 82–83). Wachal (2002, 201) notes that the increasing taboo status of ethnic

[4] *Oxford Advanced Learner's Dictionary*, see References.

slurs has also become visible in dictionaries. *OALD* s.v. *taboo word n.*, for example, defines the word as "a word that many people consider offensive or shocking, for example because it refers to sex, the body or people's race". The way usage and the insulting nature of some words are indicated in dictionaries will be the topic of Chapter 3, and Section 3.3 will focus on the treatment of ethnophaulisms in dictionaries in particular.

3. Usage in dictionaries

In the previous chapter, word meaning was discussed on a general level. As dictionaries are the focal point of this study, it is also important to discuss the way word meaning is represented in dictionaries. In addition to word meaning in general, the notion of usage in dictionaries will also be examined in this chapter. The term *usage* can be applied to all sorts of limitations that using a particular word may have (Landau 2001, 217). For example, a word may not be 'neutral' in the sense that it is offensive. Section 3.1 of this chapter will focus on word meaning and usage in dictionaries.

As the topic of the study is insulting nationality words, this chapter will also examine the ways of marking potentially offensive words in dictionaries. Section 3.2 will shed light on this matter. Furthermore, it is also necessary to study the way ethnic slurs in particular are treated in dictionaries. This issue will be handled in Section 3.3.

3.1 Indicating meaning and usage in dictionaries

Defining a word in a dictionary is arduous, and there are several problems related to this process. Dictionary definitions need to be as concise as possible, and a great deal of editing work is usually required before the definition reaches its final form (Landau 2001, 154). It is not an easy task to pin down the meaning of a complex concept with a sentence or two. Béjoint (2010, 320) states that "[t]he definition of a word is not its meaning: it is an attempt at describing its meaning in such a way that it will clarify it". Jackson (1988, 126) notes that the definition can only be seen as a potential meaning of a word, and that the final meaning will be formed when the word is uttered or written in context. In fact, both Béjoint (2010, 301) and Hanks (2000, 125) point out that numerous scholars have challenged the idea of a word having a fixed number of identifiable senses and also the whole concept of word meaning.

When definitions are composed, the compilers of a dictionary can turn to numerous sources. Corpora are of significant help, but lexicographers also use citation files. A citation file can be defined as "a selection of potential lexical units in the context of actual usage, drawn from a variety of written sources and often some spoken sources, chiefly because the context illuminates an aspect of meaning" (Landau 2001, 190). Citation files are particularly useful for historical information, as they often have information that is older than the content in corpora, and they are also useful for collecting examples of new words (ibid., pp. 191–193). Publishers have their own citation databanks, and the material is collected by both in-house members and "interested members of the public" (Jackson 2002, 166). In addition to these two sources of information, lexicographers can also ask for the help of specialists of certain fields or use reference books (including dictionaries) (Landau 2001, 213–214). What is more, sometimes the compilers use introspection, their own intuitions on usage, even though it can be questioned whether this is a reliable method (Béjoint 2010, 356).

Defining words is not the only challenge facing the editors of a dictionary. When compiling a dictionary, lexicographers must decide which words to include and what to exclude. Editors need to decide which types of specialised vocabulary they will include: whether to include words that are only used in certain regions, words that are very literary, and so forth (Atkins and Rundell 2008, 182). Not all words can be used in all contexts: several limitations often restrict the choice of words. The term *usage* can be used to denote all of these limitations that using a particular word may have (Landau 2001, 217) and this view will be adopted in the present study. Usage can also denote either "all uses of language" or "the study of good, correct, or standard uses of language, as distinguished from bad, incorrect, and nonstandard uses" (ibid.). Ptaszynski (2010, 411) points out that the terminology related to marking restricted usage in dictionaries is diverse. He (ibid.) mentions the terms *linguistic labels* (used e.g. by Atkins and Rundell 2008), *usage information* (used e.g. by Landau 2001), *diasystematic*

information, *diasystematic marking*, *restrictive labels* and *stylistic glosses*. In this study, Landau's (2001, 217) term *usage information* will be adopted.

Usage information can be marked in various ways, but it is usually stated with the help of labels in a dictionary entry (Ptaszynski 2010, 412). Atkins and Rundell (2008, 226) define linguistic labels as markers of the vocabulary type that the headword represents. The *OED2*[5] s.v. *label n.1*, draft additions b., defines *label* as "a word or phrase used (often in italics) to specify the geographical area, register, etc., to which the term being defined belongs". Labels are usually explained in the front or back matter of a dictionary (Atkins and Rundell 2008, 226).

Scholars have introduced various systems for categorising labels. Landau (2001, 217–218) lists eight types of usage information typically given in dictionaries: currency or temporality; regional or geographic variation; technical or specialised terminology; restricted or taboo sexual and scatological usage; insult; slang; style, functional variety or register; and finally, status or cultural level. Jackson (1988, 150–153), on the other hand, only lists dialect and time, formality and status, and domain. Jackson (ibid.) discusses taboo words in connection with formality and status, but usage labels indicating offensiveness are not mentioned. This is probably because people have become more sensitive to the issues of racism and sexism over the previous decades, and these issues are discussed more nowadays. In addition, labels indicating offensiveness were probably not used as extensively in the 1980's as they now are.

Atkins and Rundell (2008, 227–230) list seven categories for labels: domain, region, register, style, time, attitude and meaning type. It should be noted that they (ibid.) list offensive terms as a subset of register labels, but include labels such as *pejorative* and *derogatory* in attitude labels (see Section 5.2 for the discussion of the labels used in the dictionaries of the present study). However, in their tree diagram (ibid., p. 183), offensive words have been listed as their own vocabulary type, not

[5] *The Oxford English Dictionary*, second edition; see References.

belonging to register. Thus, even scholars seem to be uncertain about how to classify various labels and how to provide the reader with the necessary usage information.

According to Landau (2001, 233), sometimes a label is not enough; the meanings of words are complex and defining the connotation of a word may require a more comprehensive usage note as opposed to a single label. In fact, labels are not the only way to convey the usage information of a word to the reader: it can also be stated in the definitions themselves or in more comprehensive usage notes (Ptaszynski 2010, 412). Examples of different ways of indicating one type of usage information, the potential offensiveness of words, will be given in Section 3.2. Ptaszynski (ibid.) also argues that usage information can sometimes be stated in the front or back matter of the dictionary. However, he (ibid.) does not explain the way usage information concerning individual words can be expressed when using this technique.

According to Ptaszynski (2010, 413), the type of the dictionary in question and its target audience should also be taken into account when labelling headwords and giving usage information (ibid.). Norri (2000, 91) notes in his study concerning the labelling of derogatory words that learner's dictionaries tend to use more labels and usage notes than other dictionaries. According to Norri (ibid.), this is probably because the target audience of learner's dictionaries consists of non-native English speakers and for this reason the dictionary compilers need to use more cautionary labels and usage notes.

3.2 Marking insult in dictionaries

Dictionary labelling reflects the views of the whole society to some extent. Landau (2001, 232) notes that dictionaries operate under the laws and norms of a specific society, and those laws and norms are also reflected in dictionaries. Governments have started to pay more attention to minority rights and

discrimination has become criminalised, and for this reason dictionaries have to treat potentially insulting words meticulously (ibid.).

As was the case with usage information in general, offensiveness can be stated in various ways. The following examples illustrate different ways of conveying the insulting nature of a word to the reader, (3) being an example of using labels, (4) demonstrating how offensiveness can be stated in the definition and (5) as an instance of combining a usage note with a label and also stating the offensiveness of the word in the definition:

> (3) *BED*[6] s.v. *pommy n.*: "*ANZ* (*informal humorous or disapproving*) […] a British person"
> (4) *DCS*[7] s.v. *Pommy, Pommie n.*: "*Australian* (a person who is) British. The standard, and usually derogatory, slang term for natives of or immigrants from the British Isles [...]"
> (5) *LDCE*[8] s.v. *frog n.,* sense 3: "*taboo* a very offensive word for someone from France. Do not use this word."

As concerns (5), the end part of the entry, "[d]o not use this word" is treated as a usage note in this study. Often, usage notes concerning grammar, for instance, are more elaborate and isolated in their own box below the definition. However, in this study the warnings placed at the end of the definition, as is the case with (5), will be classified as usage notes (see also Subsection 6.1.1, which discusses the categorisation into usage notes even further).

Landau (2001, 218) lists numerous labels that can be applied to potentially offensive words: *offensive*, *insult*, *disparaging*, *derogatory*, *disapproving*, *contemptuous*, *sexist* and *racist*. Not all of these labels are used in every dictionary. Each dictionary uses its own set of labels and defines the labels in its own way. The labels used by the dictionaries included in this study will be introduced in greater detail in Section 5.2. Examples (6), (7) and (8) below, taken from *Random House Webster's Unabridged Dictionary* (p. xvii), illustrate how this dictionary defines the labels it uses in connection with potentially insulting words:

---

[6] *Bloomsbury English Dictionary*, see References.
[7] *Dictionary of Contemporary Slang*, see References.
[8] *Longman Dictionary of Contemporary English for Advanced Learners*, see References.

> (6) ***Disparaging*** This label indicates that a term is used with disparaging intent, as to belittle a particular racial, religious, or social group.
> (7) ***Offensive*** This label indicates that the term so labeled is likely to be perceived as offensive by a listener or reader, whether or not any offense was intended.
> (8) ***Vulgar*** Vulgar terms are considered inappropriate in many circumstances because of their association with a taboo subject. Major taboo subjects in English-speaking cultures are sex and excretion and the parts of the body associated with those functions.

As (6), (7) and (8) show, a word can be insulting in several ways. Some labels, such as *disparaging*, refer to the register and the focus is on the person who utters the insult: the purpose is to show disapproval or contempt. On the other hand, as concerns some labels, such as *offensive*, the focus is more on the one who is being insulted. As Norri (2000, 77) notes, a word may simultaneously be both offensive and disparaging. What is more, these labels may also be modified in dictionaries in order to show different levels of offensiveness: in the American edition of the *Encarta World English Dictionary* (p. xix), it is stated that "[w]ords not universally regarded as offensive but likely to give offense in various degrees are qualified accordingly: *often considered offensive*, *sometimes considered offensive*, and *offensive in some contexts*". In addition, even though the label *taboo* (or *vulgar*) is usually used when the content of the word is seen as somehow unpleasant to talk about (such as words related to excretion), these labels are sometimes used in connection with ethnic slurs and nationality words (see Wachal 2002). As Wachal (2002, 201) notes, ethnic slurs seem to have attained a taboo status in our culture.

In this study, the terms indicating negative attitude, such as *offensive* and *insulting*, will be used as 'umbrella terms' including also possible disparaging and pejorative connotations. That is to say, if it is stated that "insulting nationality words are more often found in British than American dictionaries", the word *insulting* is used to refer to words that are offensive, disparaging or both. The distinction into offensive and disparaging uses will only be made when these particular labels in dictionaries are discussed, and when examining whether a word has more often been labelled as *offensive* than *disparaging*, for example (see Section 6.2).

Problems arise because of the fact that lexicographers do not have universal criteria to tell them which words should be regarded as offensive, which as disparaging and which should not be classified as insulting at all (Landau 2001, 233). Corpora and citation files are not very useful when it comes to labelling insulting words: according to Landau (ibid., pp. 233–234), there is often not enough context to decide whether someone has been insulted or not. Furthermore, there are also numerous instances of non-offensive uses of potentially insulting words in corpora (ibid.). The in-group and out-group use was discussed in Section 2.3, where it was mentioned that a slurring term can sometimes be used positively, to strengthen the feeling of belonging. This shows that labelling a potentially insulting word is a challenging task: the label *offensive* is not always enough to illustrate the various connotations of a word.

What is more, even if an offensive word has been used in order to insult someone, the target of the insult does not always show that s/he has been offended. In fiction, the author sometimes assumes that the reader realises whether the word used has been insulting and thus the behaviour of the assaulted character does not have to emphasise the insult (Landau 2001, 233–234). This makes using corpora even more problematic when labelling offensive words.

Regional differences also affect labelling: some words are more pejorative in a certain variety of English than others (Norri 2000, 73). In addition, as Norri (ibid.) notes, "[v]ocabulary... is always in a state of flux". Words that are derogatory may develop new opposite meanings and neutral words may develop negative connotations (ibid.). The process in which a word adopts more positive connotations is called *amelioration*, and if the opposite happens, that can be referred to as *pejoration* (Harley 2006, 104).

It should also be noted that even though the world views of the whole society have an effect on labelling, labelling is also a matter of the editor's personal opinions (Landau 2001, 232). Landau (ibid., p. 234) argues that "the lexicographer is compelled to use his own experience, moderated of necessity

by his own moral views, whether consciously or not". Brewer (2005, 262) notes that even though the *OED* may seem "impersonally authoritative", the subjectivity and personality of the lexicographers is visible in labelling in particular. Thus, it is extremely difficult, or even impossible, for the lexicographer to be fully objective when labelling words; it is always a matter of personal views, too.

It is a difficult task to decide the appropriate way of indicating the potential offensiveness of a word, because there are several things that have to be taken into account. In addition to this, different dictionaries have their own policies as concerns defining and labelling insulting words. This causes even more variation among labels and other markers of usage information. As Norri (2000, 92) notes, this is natural because each dictionary is a separate work. However, there tend to be discrepancies within a single volume as well (ibid., p. 93), and this, of course, is not desirable.

3.3 Ethnophaulisms in dictionaries

During the first half of the twentieth century, ethnophaulisms[9] were not usually labelled in any way in dictionaries (Béjoint 2010, 208). However, as the general attitude towards minorities changed and hate speech became criminalised in many countries, the representation of ethnic slurs in dictionaries also changed. Wachal (2002, 201) notes that "ethnic slur terms have been getting more tabooed in dictionaries, as in our culture". As Bauer (1994, 145) mentions, language change can be imposed and deliberately sought, and this is what has happened with ethnophaulisms.

When dictionary editors started to pay more attention to potentially offensive ethnic words and nationality words, these words were first sometimes omitted from dictionaries in order to avoid offending anyone (Béjoint 2010, 207). In the 1970's, 1980's and 1990's, these words started to appear in dictionaries combined with labels such as *offensive* and *derogatory* (ibid., pp. 207–208). More complicated formulations, such as "a rude word for..." were also used to inform the reader of

[9] See Section 2.4 for the definitions of *ethnophaulism* and *ethnic slur*.

offensiveness (ibid., p. 208). Nowadays, "[e]very major English dictionary [...] takes a stand on the side of those who deplore racial and ethnic bigotry" (Landau 2001, 232). Even some specific symbols for marking ethnophaulisms have been introduced: in the seventh edition of *Concise Oxford English Dictionary* (1982), the letter *R* was used to indicate racially offensive usage (Béjoint 2010, 208).

Nevertheless, a consistent system has never been established when it comes to marking ethnic slurs (Béjoint 2010, 208). Landau (2001, 232) states that not all ethnophaulisms are treated in a similar manner in dictionaries: ethnic slurs targeting groups that have been severely abused in the past, such as African Americans, are "regarded with special outrage". Dictionaries have their own systems for marking ethnophaulisms, but, according to Béjoint (2010, 208), "whatever method is used it is never enough". The system is not coherent, and "a label is never as forceful as the presence or absence of a word or phrase" as labels are often ignored by dictionary users (ibid.) – even though one could argue that a label, too, is a word.

Another problem related to marking ethnophaulisms in dictionaries is that sometimes dictionaries use labels almost too often when trying to avoid offending anyone (Landau 2001, 234). Landau (ibid., p. 163) argues that sometimes the offensiveness of a word may be seen as a more important item than the definition itself. Landau (ibid.) mentions *Encarta World English Dictionary* as an example: many definitions in this dictionary start with a phrase such as "An offensive word for...". In those cases, the essential meaning of the word is seen as being an offensive term, and other components of meaning are seen as secondary (ibid.)

This chapter has examined the various ways in which dictionaries formulate their definitions and how they present usage information, in particular with potentially offensive words. As stated in the previous chapter and in the Introduction, ethnic slurs denoting some nationality may be regarded as more insulting than ethnophaulisms denoting some other nationality. For example, words denoting the previously-oppressed groups may be marked as more offensive than words denoting their oppressors.

The next chapter will focus on the relations between the nationalities studied in this thesis, and thus it will also shed light on the nationalities in the present selection that are potentially treated with the most special care in dictionaries.

4. Historical background: the relations between the nationalities examined

One of the themes of this study pertains to the question whether words denoting a certain nationality are marked as more insulting than words denoting some other nationality in dictionaries. Landau (2001, 232) states that ethnophaulisms targeting groups that have been severely abused in the past are "regarded with special outrage", and Norri's (2000, 83) study also points out that words denoting oppressed groups seem to be marked as more offensive than those denoting the non-oppressed. As Loomba (1998, 7) notes, it is "debatable whether once-colonised countries can be seen as properly 'post-colonial'". Therefore, it is useful to examine the relations between the nationalities included in the study: it will shed light on the nationalities that can be regarded as 'the oppressors' and those that may be considered 'the oppressed'.

It was stated in the previous chapters that not all nationalities are evaluated in a similar manner: some nationalities are seen in a more positive light than others, and some ethnophaulisms appear to be treated with greater discretion than others in dictionaries. An interesting juxtaposition seems to rise from these arguments: on the one hand, it is argued that ethnic groups that have been abused and oppressed in the past are treated with special care in dictionaries (see Landau 2001, 232, for example). Battistella (2005, 83) notes that "[r]elatively unassimilated groups are more likely to be targets of majority xenophobia than groups which are more assimilated into the dominant culture". It is often the case that members of the oppressed or abused groups may have a hard time assimilating into the dominant culture, when entering a new country as a refugee or an asylum seeker, for instance. On the other hand, Rice et al. (2010, 121) argue in their study that sometimes people consider insults targeting their own group more insulting than insults targeting some other group. Therefore, it will be interesting to see which nationality words will be treated as the most and which as the least offensive ones in dictionaries.

This chapter focuses on the history of the most recent centuries (including the present-day situation), as the dictionaries studied are modern editions and the events that have taken place several hundreds or thousands of years ago are not likely to have as significant an effect on the attitudes of today as more recent events. The dictionaries included in the study are British and American volumes, and thus the relations between the nationalities studied will be examined from the point of view of Great Britain and the United States.

The concept of *Britishness* needs to be shortly discussed before discussing Britain's relations to the other nationalities studied. As the majority of the dictionaries used in the study are published in Britain (see Section 5.1), it was natural to choose the British as one of the nationalities that would be examined. However, referring to the Welsh, the Scottish, and the English as *the British* is not entirely unproblematic. As Julios (2008, 3) points out, the "British identity" has long been and still is a debated concept. For instance, the nationalism in Scotland has questioned the existence of the British identity. As Leith and Soule (2011, 139) note, the Scottish have for a while been promoting the idea of Scotland as a nation as opposed to "simply a 'sub-national' component of the UK". In the fall of 2014, a referendum on Scottish independence took place, but Scots voted against becoming an independent country by 55% to 45% ("Scottish referendum: Scotland votes 'No' to independence", BBC online, 19 September 2014). Nonetheless, in this study words denoting the Scottish, the Welsh and the English (and words denoting all Brits) will be grouped together under the nationality of *the British* as the focus is not only on Britain (see also Section 5.4 for the meanings of the words examined).

Britain shares a great amount of history with all the nationalities studied (with the exception of Mexicans). Even though the British have had colonies in the United States and oppressed the native Americans, for instance, the Americans cannot be assigned the role of the oppressed as the United States is one of the superpowers of today. Even though fighting against each other in the American Revolutionary War, the British and the Americans have fought as allies afterwards, including the

Second World War. Neither can the British be seen as oppressors of the French, Germans or Mexicans: even though Brits have several times been at war against both the French and Germans, a situation where the British would clearly have been the oppressors (or the oppressed) for a long period of time has not occurred in recent history with the French or Germans. Moreover, Great Britain and Mexico do not share very much common history, which is natural due to the geographical distance. For this reason, it is also understandable that the roles of the oppressors and the oppressed cannot be assigned in this situation.

The nationality that the British can be seen to have oppressed are the Irish. A long period of their common history is characterised by Britain's colonialism over Ireland. Therefore, Ireland can be seen to belong to 'the oppressed', and Britain to 'the oppressors'. As Coakley (2012, 189) notes, "[t]he incorporation of Ireland into the extended English/British state took the form of a full-scale conquest, involving massive human and physical destruction". Thus, one would expect the words for the Irish to be treated as more insulting in British dictionaries than words for the British. Nevertheless, this may not be the case as sometimes people consider insults targeting their own group more insulting than insults targeting some other group (Rice et al. 2010, 121).

As far as the Americans are concerned, it was already stated that roles of the oppressor and the oppressed cannot really be assigned to their situation with the British. Nor do these roles seem to apply to America's relations to the French, Germans or the Irish, either way. The relations between the United States and Mexico, on the other hand, are different. Due to geographical proximity, these two nations interact regularly. The relations between the United States and Mexico in previous centuries will not be gone into here. The tension in the present-day relations between these two countries is likely to have a greater effect on the dictionary policies than the issues of the earlier centuries. The problem of (illegal) immigration from Mexico to the United States has been discussed extensively over the last few decades. As Aguila et al. (2012, 2) argue, Americans and Mexicans see the situation

differently: in Mexico, migration to their northern neighbour is seen as a way of improving one's life quality, whereas in the United States Mexicans are sometimes seen as a threat to national security and identity as well as causing harm to public finances. It will be interesting to examine whether this situation has an effect on the dictionary policies; whether the words for Mexicans will be treated with special care in order to avoid the impression that the dictionary compilers would regard Mexicans as a threat, for example. In addition, some people may see the United States as the oppressor of some other nationalities too, due to its role as the leading superpower of today.

5. Materials and methods

In this chapter, the materials and methods used in the present study will be introduced. The material of the study consists of twenty dictionaries and a corpus. The dictionaries examined will be discussed in Section 5.1. They will be categorised into different types and each dictionary used in the study will be listed. The labels used by these dictionaries for offensive words vary from one dictionary to another, and for this reason it is useful to list all the labels used by the dictionaries along with their definitions. The labels with their definitions will be listed in tabular form in Section 5.2. In Section 5.3, the British National Corpus —the corpus used in the study— will be introduced.

In total, 37 potentially offensive nationality words will be examined in this thesis. The words will be categorised into subgroups and these groups will be presented in Section 5.4. Furthermore, definitions of each word examined will also be given in Section 5.4. The methods of the lexical analysis will be outlined in Section 5.5, comprising methods for both the dictionary and corpus analysis. The potential complexities relating to the study will also be introduced in Section 5.5.

5.1 Dictionaries studied

The main material for this study is various monolingual dictionaries. Dictionaries can be classified in several ways. Landau (2001, 7) introduces Malkiel's system in which dictionaries are classified according to their range, perspective and presentation. Béjoint (2010), on the other hand, distinguishes dozens of types according to publishing place (British and American dictionaries), target audience (learner's dictionaries, for example), coverage (general purpose and specialised dictionaries) and many more.

Jackson (1988, 159) suggests that dictionaries can be divided into *general purpose dictionaries* and *specialist dictionaries*. *General purpose dictionaries*, as the name suggests, aim at giving a comprehensive coverage of the whole vocabulary, and are not restricted to a particular vocabulary type

or target audience (ibid.). *Specialist dictionaries*, on the other hand, are more restricted. They can be further categorised into two groups: those providing specialist information, such as a pronunciation or medical dictionary, and those aimed at a special target audience, such as a children's dictionary (ibid., p. 165).

In this study, I will divide the dictionaries into three categories: general purpose dictionaries, learner's dictionaries (including collegiate dictionaries) and slang dictionaries. Thus, they represent the types mentioned by Jackson (1988, 159–165), learner's dictionaries belonging to specialist dictionaries aimed at a special audience and slang dictionaries representing dictionaries that provide specialist information. In addition, general purpose dictionaries and learner's dictionaries will further be divided into British and American volumes in this study.

It should be noted that using the term *British* in connection with the other half of the dictionaries is not entirely straightforward, as most of the dictionaries are published in England. However, some of the works studied are published in Scotland, such as *Collins English Dictionary* (see References), and thus *British* will be used to include both England and Scotland as the place of publication. Furthermore, the division into British and American volumes is a common practice among scholars examining dictionaries, and this system is used by Béjoint (2010), for example.

The three-part classification will be enough as the material consists of twenty dictionaries. This number will be sufficient to offer a comprehensive picture of the words studied. There are more British dictionaries than American ones as the emphasis will be more on Britain. In a study of this scope, the focus has to be narrowed down. It will be possible to include more dictionaries (and nationalities) in further research.

The general purpose dictionaries included in the study are large volumes, so called desk-dictionaries (see Jackson 1988, 160), and they include more words than learner's dictionaries. Some of the words included in this study are quite rare, and for this reason half of the dictionaries that are

included in the analysis are general purpose dictionaries. The British general purpose dictionaries that will be consulted are *Bloomsbury English Dictionary* (hereafter *BED*), *The Chambers Dictionary* (*ChD*), *Collins English Dictionary* (*CED*), the British edition of *Encarta World English Dictionary* (*EWEDL*), *Oxford Dictionary of English* (*ODE*) and the online version of *The Oxford English Dictionary* (the *OED*). It should be noted that the *OED* is not exactly a general purpose dictionary but a historical one, but as the etymology of words is not in the focus of this study the *OED* will be classified among the general purpose dictionaries due to its comprehensiveness. Furthermore, as the *OED* is currently being updated into the third edition, some entries are already updated and some are still being processed. The abbreviation *the OED2* will be used to refer to the entries which are not yet fully updated, and, correspondingly, the abbreviation *the OED3* will be used to refer to the entries which have already been updated and belong to the third edition of the *OED*. As far as other dictionaries are concerned, only one edition has been used.

In addition, the following American general purpose dictionaries are included: *The American Heritage Dictionary of the English Language* (*AHD*), the American edition of *Encarta World English Dictionary* (*EWEDNY*), *Random House Webster's Unabridged Dictionary* (*RHD*) and *Webster's Third New International Dictionary* (*W3*). Both *RHD* and *W3* include a section for new words added to these particular editions (*Addenda* in *W3* and *New Words Section* in *RHD*). It will be specified if the entry examined belongs to these additional sections.

Even though learner's dictionaries are not likely to include many of the words studied, they will also be consulted because learner's dictionaries tend to provide more warnings when it comes to derogatory words (Norri 2000, 91). I will examine whether this is also the case with the nationality words studied. In this MA thesis, I will consult the following British learner's dictionaries: *Cambridge Advanced Learner's Dictionary* (*CALD*), *Collins COBUILD Advanced Learner's Dictionary* (*COBUILD*), *Longman Dictionary of Contemporary English for Advanced Learners* (*LDCE*), the

online version of *Macmillan English Dictionary* (*MED*) and *Oxford Advanced Learner's Dictionary* (*OALD*).

The United States does not have a similar tradition for learner's dictionaries as Britain. In this study, collegiate dictionaries will be treated as an American counterpart for British learner's dictionaries, even though they are not exactly similar. According to Béjoint (2010, 190), British learner's dictionaries are usually targeted towards EFL (English as a foreign language) students, whereas American ones are mostly designed for ESL (English as a second language) students. Jackson (2002, 67) adds that collegiate dictionaries are also aimed at high school and undergraduate students, in addition to ESL students. In the preface of *The American Heritage College Dictionary* (see References), p. iv, it is stated that the dictionary editors' aim has been "to provide a dictionary of the highest quality at an affordable price for college students – indeed, for anyone who has a need or desire to learn more about words", and that the dictionary aims at using a clear defining style with a great number of examples. In the preface of *Merriam-Webster's Collegiate Dictionary* (see References), p. 6a, it is mentioned that even though the title "may suggest a special appropriateness for the older student", its target audience is not only limited to students. The blurb on the book jacket of *Webster's New World College Dictionary* (see References) mentions that the dictionary "is a practical, authoritative resource that is perfectly suited for home, school, or the office". Thus, the collegiate dictionaries themselves mention students as one of their target groups, but emphasise the fact that other user groups will benefit from these dictionaries, too. The collegiate dictionaries that will be examined in this study are the aforementioned *The American Heritage College Dictionary* (*AHCD*), *Merriam-Webster's Collegiate Dictionary* (*MWCD*) and *Webster's New World College Dictionary* (*WNW*).

Many of the words included in this study represent slang and thus it is natural to also include some slang dictionaries. Furthermore, slang dictionaries have not been included in the study by Norri (2000), for example, and therefore it will be interesting to see whether their policies of marking potentially

offensive words differ from the two other types of dictionaries. The slang dictionaries that I will investigate are *Dictionary of Contemporary Slang* (*DCS*) and *The Concise New Partridge Dictionary of Slang and Unconventional English* (*DSUE*).

I have aimed at using the latest possible editions in my analysis of the dictionaries. However, in some cases the newest editions have not been available and in those cases the second newest editions have been used. Furthermore, in some cases only parts of the dictionary have been updated; *W3* and *RHD*, as was mentioned above, both have an updated section. For this reason, some entries are significantly older than others. On the other hand, it is interesting to see whether there are differences between newer and older dictionaries. I assume that newer volumes use more labels and warnings when it comes to insulting nationality words.

5.2 Labels in the dictionaries studied

Different dictionaries use different labels when presenting usage information. Labelling is an important way of marking offensiveness in a dictionary. Each dictionary uses its own set of labels and the definitions of these labels vary. For this reason, it is useful to sum up the labels used in the dictionaries studied. The labels used with potentially offensive and disparaging words with their definitions are listed in Table 1 below, the labels and the dictionaries being listed in an alphabetical order:

Table 1. Labels for insulting words in the dictionaries studied

Label	Definition (dictionary)
Derogatory	"This label is used for words that have meanings which are not in themselves negative or disparaging, yet the words invariably convey negative or disparaging judgments. It may appear alone or in combination as *Derogatory Slang.*" (*AHD*, p. xxiii)

Label	Definition (dictionary)
Derogatory (cont.)	"implies that the connotations of a word are unpleasant with intent on the part of the speaker or writer." (*CED*, p. xvi) "language intended to convey a low opinion or cause personal offence." (*ODE*, p. xiv)
Disapproval	"The label DISAPPROVAL indicates that you use the word or expression to show that you dislike the person or thing you are talking about. An example of a word with this label is *infantile*" (*COBUILD*, p. xv)
Disapproving	"marks a derogatory attitude on the part of the speaker" (*BED*, p. xviii; *EWEDL*, p. xviii; *EWEDNY*, p. xix) "used to express dislike or disagreement with someone or something" (*CALD*, p. i) "expressions show that you feel disapproval or contempt, for example *blinkered, newfangled*" (*OALD*, p. i)
Disparaging	"This label indicates that a term is used with disparaging intent, as to belittle a particular racial, religious, or social group." (*RHD*, p. xvii)
Impolite	"not taboo but will certainly offend some people" (*MED*, "Labels and abbreviations")
Insult	"a pejorative term that would be likely to insult or upset somebody if (it is) said directly to the person" (*BED*, p. xviii; *EWEDL*, p. xviii; *EWEDNY*, p. xix)
Not polite	"a word or phrase that is considered rude, and that might offend some people" (*LDCE*, p. i)
Offensive	"This label is reserved for words and expressions such as racial, ethnic, or gender slurs that are not only derogatory and insulting to the person to whom they directed but also a discredit to the one using them. This label may occur alone or in combination as *Offensive Slang*." (*AHCD*, p. xvi) "This label is reserved for words and expressions such as racial, ethnic, or gender slurs that are derogatory and insulting to the members of the group to whom they are directed. This label may occur alone or in combination as *Offensive Slang*." (*AHD*, p. xxiii) "likely to be offensive to many people, for example because of being racist or sexual" (*BED*, p. xviii) "very rude and likely to offend people" (*CALD*, p. i) "indicates that a word might be regarded as offensive by the person described or referred to, even if the speaker uses the word without any malicious intention." (*CED*, p. xvi)

Label	Definition (dictionary)
Offensive (cont.)	"likely to offend people, or to insult them; words labelled offensive should therefore usually be avoided, e.g. **cripple**" (*COBUILD*, p. xiv) "likely to be offensive to many people, for example, because it is racist or sexual" (*EWEDL*, p. xviii; *EWEDNY*, p. xix) "extremely rude and likely to cause offence" (*MED*, "Labels and abbreviations") "expressions are used by some people to address or refer to people in a way that is very insulting, especially in connection with their race, religion, sex or disabilities, for example *half-caste, slut*. You should not use these words." (*OALD*, p. i) "language that is likely to cause offence, particularly racial offence, whether the speaker intends it or not." (*ODE*, p. xiv) "This label indicates that the term so labeled is likely to be perceived as offensive by a listener or reader, whether or not any offense was intended." (*RHD*, p. xvii)
Rude	"used mainly to describe words which could be considered taboo by some people; words labelled rude should therefore usually be avoided, e.g. **bloody**" (*COBUILD*, p. xiv)
Showing disapproval	"used when it is not obvious from a definition that a word says something bad about someone or something: *babyish*, *self-satisfied*" (*MED*, "Labels and abbreviations")
Taboo	"marks classic taboo words referring to sex and bodily functions" (*BED*, p. xviii) "indicates words that are not acceptable in polite use." (*CED*, p. xvi) "for classic taboo words referring to sex and bodily functions" (*EWEDL*, p. xviii; *EWEDNY*, p. xix) "a word that should not be used because it is very rude or offensive" (*LDCE*, p. i) "expressions are likely to be thought by many people to be obscene or shocking. You should not use them. Examples are *bloody, shit*." (*OALD*, p. i)
Very offensive	"highly likely to offend people, or to insult them; words labelled VERY OFFENSIVE should be avoided, e.g. **wog**" (*COBUILD*, p. xiv)
Very rude	"used mainly to describe words which most people consider taboo; words labelled VERY RUDE should be avoided, e.g. **fuck**" (*COBUILD*, p. xiv)
Vulgar	"This label warns of social taboos attached to a word; it may appear alone or in combination as *Vulgar Slang*" (*AHCD*, p. xvi)

Label	Definition (dictionary)
Vulgar (cont.)	"This label warns of social taboos attached to a word; it may appear alone or in combination as *Vulgar Slang*, which is used for words that violate accepted standards of decency." (*AHD*, p. xxiii) "Vulgar terms are considered inappropriate in many circumstances because of their association with a taboo subject. Major taboo subjects in English-speaking cultures are sex and excretion and the parts of the body associated with those functions." (*RHD*, p. xvii) "The word or meaning is regarded by many people as being indecent or extremely coarse and hence unsuitable for use in many social situations." (*WNW*, p. xx)
Vulgar slang	"informal language that may cause offence, often because it refers to the bodily functions of sexual activity or excretion, which are still widely regarded as taboo" (*ODE*, p. xiv)

As can be seen from Table 1, some of the dictionaries mentioned in Section 5.1 are absent from the table. Not all dictionaries examined in the study, such as *DCS*, use labels to mark the offensiveness of a word. *DCS* comments on offensiveness in the definitions, whereas *W3* and *WNW* employ usage notes, with the exception of *WNW* introducing the usage label *vulgar* (see *W3*, p. 18a, and *WNW*, p. xx). Some dictionaries, on the other hand, use labels but do not provide the reader with definitions. *ChD* (pp. xxii–xxiv), for example, only lists the abbreviations used in the dictionary with their full forms, but does not give any definitions. The only abbreviation used with offensive words in this list is *derog* for *derogatory* (ibid.). It is stated in *ChD* (p. xxii) that "many of the abbreviations are used as labels, but other, unabbreviated, labels are also found in the dictionary, all of which should be self-explanatory". It can be questioned whether these labels are self-explanatory, because Table 1 above shows that the definitions of the labels used vary significantly. The policy of *DSUE* (p. xxxiii) is similar to that of *ChD*: it only lists the abbreviations used in the dictionary with their full forms, the relevant ones being *derog.* for *derogatory*, *pej.* for *pejorative(ly)* and *vulg.* for *vulgarism*. In addition, the *OED* does not provide the users with definitions of the labels used, either; again, only the

abbreviations with their full forms are listed (the *OED*, "Abbreviations"). Furthermore, *MWCD* (p. 19a) lists four stylistic labels that can be used with offensive words, but does not provide the reader with a definition of any single label. Instead, they are defined collectively: "[t]he stylistic labels *disparaging*, *offensive*, *obscene*, and *vulgar* are used for those words or senses that in common use are intended to hurt or shock or that are likely to give offense even when they are used without such an intent" (*MWCD*, p. 19a).

Table 1 shows that there is a great deal of variation among the labels and their definitions. Many labels are only used by a single volume, such as *rude* (used only by *COBUILD*) and *disparaging* (listed only in *RHD*). The label *offensive* is the most common one as it is found in twelve of the dictionaries examined. The definition, however, varies: according to *ODE* (p. xiv), this label is particularly used in connection with racial insults, whereas some dictionaries, such as *CALD* (p. i) and *MED* ("Labels and abbreviations") do not specify any group of words relating to this particular label.

Even though the dictionaries examined do not treat labels unanimously, three broad categories can be distinguished based on the listings of Table 1. These groupings are based on the way the dictionaries themselves explain their labels. Firstly, there is a difference according to whether a word or expression is seen as offensive on the part of the addresser or the addressee. The labels *derogatory*, *disapproval*, *disapproving* and *disparaging* indicate that the speaker or writer using the offensive expression intends to insult the addressee. *Showing disapproval* is not as easy to categorise as the previously mentioned labels, but it could also be placed under this category: the definition of the word itself may not be offensive, but it can still be used with a belittling intent. With the labels *impolite*, *insult*, *not polite*, *offensive* and *very offensive*, on the other hand, the focus is more on the addressee who might be offended even if no offence was intended (even though *insult* seems to combine both the disparaging and offensive connotations). A third category which can be derived from Table 1 is that of taboo subjects, including the labels *rude*, *taboo*, *very rude*, *vulgar* and *vulgar slang*. Mostly these

taboo labels are used with words denoting sex or bodily functions, but according to some dictionaries, they can be extended to denote other subjects, too. According to *LDCE* (p. i), for example, the label *taboo* can be applied to any word which is seen as rude or offensive. This is why such labels are included in the table, as they may sometimes occur with ethnophaulisms. In addition, it was stated in previous chapters that ethnophaulisms are nowadays often seen as taboos.

These three categories presented above may sometimes overlap: the above-mentioned example of the definition of *taboo* in *LDCE* is a case in point. “[A] word that should not be used because it is very rude or offensive” could as well be placed under the group where the reaction of the addressee is in focus. Furthermore, sometimes a word can simultaneously be used with a belittling intent and also be offensive to the addressee. This is seen in the label *insult*, for instance: it is defined as “a pejorative term that would be likely to insult or upset somebody if (it is) said directly to the person” (*BED*, p. xviii; *EWEDL*, p. xviii; *EWEDNY*, p. xix). Thus, it is simultaneously pejorative and offensive: it is used with an insulting intent on behalf of the speaker or writer, and the addressee is also likely to be offended.

As there is a great amount of variation in the labels used by the dictionaries examined, the different labels used will not be specified in the analysis in each case. Each dictionary has its own definitions of the labels, and they are not comparable in every case, as the definitions may be significantly different. The focus in the analysis of labels will be more on whether they have been used at all, and whether they have been modified with words such as *often* or *sometimes*. This method helps to distinguish the words which are treated as being more insulting than others. Furthermore, the application of three different kinds of labels —offensive, disparaging, and taboo— will be examined. The methods of the study will be presented in more detail in Section 5.5.

5.3 The corpus used

The corpus used in this study is the British National Corpus, hereafter the BNC. The BNC is a 100 million word corpus, and it represents British English of the late 20th century (Burnard 2010). The compilation of the corpus started in 1991 and it was completed in 1994 (ibid.). The edition used in this study is the CQP-edition, Version 4.3 (updated in January 2010). The BNC includes both written and spoken English, written English comprising 90% and spoken English 10% of the corpus (ibid.). The written part consists of a variety of genres: it includes newspapers, popular fiction, school essays and academic books, for example (ibid.). The spoken part also includes a variety of data, such as informal conversations, radio shows and formal business meetings (ibid.). Both parts will be used in the present study.

A British corpus was chosen because the majority of the dictionaries studied are British dictionaries. An analysis of an American corpus could be added in a further study, but in a thesis of this scope the selection of corpora had to be narrowed down to one. Furthermore, the BNC represents a wide variety of different genres, and it is a large corpus. Some of the words included in the study are quite rare, and they are more likely to be found in a large corpus than in a smaller one. In addition, the BNC represents modern British English of the later part of the 20th century. It corresponds quite well to the time range of the dictionaries examined, even though the majority of the dictionaries were published at the beginning of the 21st century. Thus, it should be noted that the examples in the BNC represent slightly older usage, which might have a minor effect on the results: being politically correct when referring to different nationalities, for example, was not as important some decades ago as it is now. What is more, it should be noted that as the BNC represents British English, words belonging to some other dialects of English are not as often found in the corpus.

5.4 Words examined

The 37 words examined were collected with the help of thesauri (*Historical Thesaurus of the Oxford English Dictionary*[10] (2009) and *Roget's Thesaurus*[11] (2002)) and scholarly articles (e.g. Rice et al. 2010 and Campbell 2006). In addition, *The Racial Slur Database* was used to find more recent words. Some of the words listed in these sources are either obsolete or very rare. For instance, Lillo (2001, 340–346) lists approximately 100 ethnic slurs, all of which are instances of rhyming slang. Many of those words are not very common, and they would not likely be listed in dictionaries. I chose words that are fairly common and still in use, so that they will be included in at least some of the dictionaries. Owing to the scope of the present study, a maximum number of 6 nationalities and 40 words in total was set when choosing the words examined. Lists for different possibilities were made, and these words were checked in a few dictionaries (the *OED*, *W3*, *EWEDL* and *DSUE*) in order to see whether they were listed at all and whether they were said to be insulting. Some of the words and their negative connotations were also priorly familiar to me. In the end, six nationalities and 37 words were left.

In this thesis, words for the British, Americans, the Irish, the French, Germans and Mexicans will be examined. These are all nationalities which use English as their national language or they are nationalities that are or have been in contact with the English-speaking world, because of wars, colonialism or some other factor. The contact with the English-speaking world is essential because the dictionaries that will be studied are monolingual English dictionaries. There are more likely to be several words for those nationalities in the English language that have been in close contact with English-speaking nations.

---

[10] The thesaurus sections examined were **01.02.07.08.06.01 (*n.*)** *British nation*, **01.02.07.08.06.01.01 (*n.*)** *English nation*, **01.02.07.08.06.01.02 (*n.*)** *Welsh nation*, **01.02.07.08.06.01.03 (*n.*)** *Scottish nation*, **01.02.07.08.06.02 (*n.*)** *The Irish*, **01.02.07.08.06.07 (*n.*)** *Native/inhabitant Germany*, **01.02.07.08.06.23 (*n.*)** *French nation*, **01.02.07.08.10.01 (*n.*)** *Native/inhabitant Australia*, **01.02.07.08.15.01.01 (*n.*)** *Native/inhabitant U.S.A.* and **01.02.07.08.15.02 (*n.*)** *Native/inhabitant Latin America*.

[11] The thesaurus section examined was **191 Inhabitant**.

It should be noted that the words examined, referring to a particular nationality, are not completely synonymous. In fact, exact synonyms are very rare, if they exist at all (Saeed 2009, 65). The connotation and even the denotation of the words examined may be slightly different even if they are placed under the same nationality group, and the descriptive and non-descriptive meaning of the words may vary (see Section 2.1 for the discussion of descriptive and non-descriptive meaning and Section 2.2 for denotation and connotation). Thus, even though two words refer to the same nationality, they may belong to different dialects, for instance. The basic meaning of the words may also be slightly different; *wetback*, for example, refers mostly to a Mexican who has entered the United States illegally "by wading or swimming the Rio Grande" (*ChD* s.v. *wetback n*., sense 1), whereas *spic* refers to a Spanish-speaking person and can thus also refer to a Puerto Rican, for instance (*OALD* s.v. *spic n*.). The differences in meaning between the words studied will be illustrated in Table 2 (see pages 47–49), which lists definitions for each of the words examined. The definitions given in the table have been composed by closely examining the dictionaries studied. Only the definitions that refer to the nationality itself have been taken into account and other senses have been excluded. The main sources used for the definitions (the entries that best represent the definitions for all the dictionaries studied) are mentioned in brackets after the definitions.

It should also be noted that the capitalisation of the words examined varies in different dictionaries. In this study, the capitalisation of the *OED* will be followed for the sake of consistency, as the *OED* lists the greatest number of the words examined, 35 out of 37 words. The two words not included in the *OED*, *seppo* and *taco bender*, were both listed in *DSUE*, and with these cases the words will be capitalised according to the way shown in *DSUE*.

The dictionaries studied are British and American volumes (see Section 5.1) and therefore Great Britain and the United States are the starting point for the study, to which the treatment of other nationalities will be compared (see Section 5.5 for the methods employed in the study). The greater

emphasis is on Great Britain as the majority of the dictionaries used in the study are British. Words for the British included in the study are *Jock*, *Limey*, *Pom*, *Pommy/-ie*, *rosbif/-beef*, *sawney* and *Taff(y)*. It should be noted that not all of these words refer to the British in general, as *Taff(y)*, for example, denotes a Welsh person (see Table 2 below). *Sweaty* (referring to a Scot, *DSUE* s.v. *sweaty n.*) was also looked up in the dictionaries, but it was only found in *DSUE*. Words for the Americans which will be examined are *seppo*, *septic (tank)*, *Yank*, *Yankee* and *Yanqui*. It should be noted that in this thesis, the word *American* will be used to denote only people from the United States, and not from Canada, for instance. In addition, *Merkin* was checked in the dictionaries but it was only found in the *OED3* (s.v. *Merkin n.2*) and not stated to be offensive. Words referring to the British and the Americans will be referred to as "Group 1".

Group 2 consists of words for the Irish. The Irish have been a part of the British empire and later gained their independence. Words for the Irish in this study are *bogtrotter*, *harp*, *Irisher*, *Mick*, *paddy* and *Pat*. In addition, *bog-lander*, *Patess*, *Patlander* and *potato-eater* were checked in the dictionaries but they were all listed only in the *OED* (the *OED2* s.v. *bog n.1*, C4; the *OED3* s.v *Patess n.*; the *OED3* s.v. *Patlander n.* and the *OED3* s.v. *potato n.*, C2). Words for Australians were also supposed to belong to this group as Australia belonged to Britain in the past, but offensive words denoting Australians were not commonly listed in the dictionaries (*kangaroo*, *boon* and *booner*, for example, were checked). The Americans, on the other hand, could be placed under this group as Britain had colonies in the United States in the past. However, words denoting Americans will be under Group 1 as many of the dictionaries studied are American volumes and thus they form the starting point with Britain. In addition, the emphasis is more on the present-day situation and the United States has had its independence for a longer period of time than Ireland.

Group 3 consists of words for the French, Germans and Mexicans. France, Germany and Mexico have all been in close contact with Great Britain and/or the United States, but they have not officially

belonged to Great Britain or USA. The words for the French are *Frencher*, *Frenchy/-ie*, *frog*, *frog-eater*, *froggy/-ie*, *mounseer* and *parleyvoo*. As far as the Germans are concerned, the following words will be analysed: *Boche*, *Fritz*, *Heinie*, *Jerry*, *Kraut* and *sausage*. *Pretzel* and *sauerkraut* were also checked but they were only listed in the *OED* (the *OED3* s.v. *pretzel n.*, sense 2, and the *OED2* s.v. *sauerkraut n.*, sense 2). The words for Mexicans examined in the study are *beaner*, *greaseball*, *greaser*, *spic/spi(c)k*, *taco bender* and *wetback*.

The words examined together with their definitions are summarised in Table 2 below. The definitions of each word in the twenty dictionaries have been examined, and all senses relating to the nationality in question are included. In Table 2, the dictionaries mentioned in brackets represent the dictionaries based on which the definitions in the table have been formed, as these dictionaries were the ones to best represent all the meanings related to a particular nationality.

It should be noted that if the entry in question mentions both the whole nationality and a subgroup of the nationality as the denotation of the word, both of these senses are listed in Table 2. However, if the sense listed is only a small subgroup of the nationality, that entry has not been taken into account in the analysis (see the following Section 5.5 on the methods of the study and on the policies of the inclusion and exclusion of particular entries and definitions, with examples). In Table 2 below, the different senses are organised in such a way that they move from general to more specific.

Table 2. Definitions of the words examined

Group	Nationality	Word	Definition
1	The British	*Jock*	A Scottish person, esp. a man; a Scottish soldier; a Scottish sailor (the *OED2* s.v. *Jock n.1*, sense 1b; *BED* s.v. *Jock n.*).
		Limey	A British person; an Englishman or an English person; a British sailor (the *OED2* s.v. *Limey n.*, sense b).
		Pom	An English or a British person in general; an immigrant from Britain (*ChD* s.v. *pom/Pom n.1*; *EWEDNY* s.v. *pom n.*; *DSUE* s.v. *Pom n.*).

Group	Nationality	Word	Definition
1 (cont.)	The British (cont.)	*Pommy/-ie*	An English or a British person in general; an immigrant from Britain (*ChD* s.v. *pommy n.*, senses 1 and 2).
		rosbif/-beef	A term used by the French to refer to an English or a British person (*ChD* s.v. *rosbif n.*; *CED* s.v. *rosbif n.*).
		sawney	A Scotsman (*W3* s.v. *sawney n.2*).
		Taff(y)	A person from Wales; a Welshman (*OALD* s.v. *Taffy n.*; *CALD* s.v. *Taffy n.*).
	Americans	*seppo*	An American (*CED* s.v. *Seppo n.*).
		septic (tank)	An American (*DCS* s.v. *septic n.*).
		Yank	An American; an American soldier (*ODE* s.v. *Yank n.*, sense 1; *WNW* s.v. *Yank n.*).
		Yankee	An American; an inhabitant of New England or the Northern United States (*AHD* s.v. *Yankee n.*, sense 3; *ChD* s.v. *Yankee n.1*, sense 1).
		Yanqui	A term used in Latin America or by Spanish-speaking Americans to refer to an American (as distinguished from a Latin American) (*MWCD* s.v. *yanqui n.*; *EWEDL* s.v. *yanqui n.*).
2	The Irish	*bogtrotter*	An Irish person; an Irish-American person (*EWEDNY* s.v. *bogtrotter n.*).
		harp	An Irish person; an Irish-American person (*DSUE* s.v. *harp n.*, sense 2).
		Irisher	An Irish person; a person of Irish descent (the *OED3* s.v. *Irisher n.*).
		Mick	A person of Irish birth or descent; an Irish-American (*RHD* s.v. *Mick n.*; *DSUE* s.v. *mick n.*, sense 1).
		paddy	A person, especially a man, of Irish birth or descent (*AHCD* s.v. *Paddy n.*).
		Pat	An Irish person; an Irishman (the *OED3* s.v. *Pat n.3*; *ChD* s.v. *Pat n.*).
3	The French	*Frencher*	A French person; a Frenchman (the *OED3* s.v. *Frencher n.*).
		Frenchy/-ie	A French person; a person of French descent; a French Canadian (*DSUE* s.v. *Frenchy n.*, sense 1; *RHD* s.v. *Frenchy n.*, sense 1).
		frog	A French person; a person of French descent (*RHD* s.v. *frog n.1*, sense 4).
		frog-eater	A French person; a person of French descent; a French Canadian (the *OED3* s.v. *frog n.1* and *adj.*, C2a; *DSUE* s.v. *frogeater n.*).
		froggy/-ie	A French person; a person of French descent; a French Canadian (the *OED3* s.v. *froggy n.* and *adj.2*, sense A2).
		mounseer	A Frenchman (*ChD* s.v *mounseer n.*, sense 2).
		parleyvoo	A French person; a Frenchman (the *OED3* s.v. *parleyvoo n.* and *adj.*, sense A2; *CED* s.v. *parleyvoo n.*, sense 3).

Group	Nationality	Word	Definition
3 (cont.)	Germans	*Boche*	A German; a German soldier (in World War I or II) (*AHD* s.v. *Boche n.*).
		Fritz	A German; a German soldier (*RHD* s.v. *Fritz n.*, sense 1).
		Heinie	A German; a German soldier (the *OED3* s.v. *Heinie n.1*).
		Jerry	A German; a German soldier (*CED* s.v. *Jerry n.*, sense 1).
		Kraut	A German; a person of German descent; a German soldier (the *OED3* s.v. *kraut n.* and *adj*., sense A2a; *AHD* s.v. *kraut n.*, sense 2).
		sausage	A German (the *OED2* s.v. *sausage n.*, sense 2d).
	Mexicans	*beaner*	A Mexican; a person of Mexican descent; a Mexican American (*ODE* s.v. *beaner n.*; *DCS* s.v. *beaner n.*).
		greaseball	A person of Latin American (esp. Mexican) or Mediterranean origin (*BED* s.v. *greaseball n.*, sense 2).
		greaser	A person of Latin American (esp. Mexican) or Mediterranean origin (*EWEDNY* s.v. *greaseball n.*, sense 2).
		spic/spi(c)k	A Spanish-speaking person, esp. a Mexican (*ODE* s.v. *spic n.*).
		taco bender	A person of Hispanic origin; a Mexican; a Mexican American (*DCS* s.v. *taco-bender n.*; *DSUE* s.v. *taco bender n.*).
		wetback	A person of Mexican birth or descent, esp. somebody who has entered the U.S. illegally (*AHD* s.v. *wetback n.*; *EWEDNY* s.v. *wetback n.*).

As Table 2 shows, the definitions of words belonging to the same nationality group vary and thus their descriptive meaning and denotation is not always the same. This is natural because absolute synonymy is rare (Cann 2011, 461). However, all words in the same nationality group also share at least some parts of their denotations. It should also be taken into account that this table does not list the different dialects that the words may belong to, for example, so the words may not be used in the same areas of the English-speaking world.

5.5 Methods employed

The first phase of the study consisted of collecting the necessary information from dictionaries. The definitions and labels used with the words studied were scrutinised and collected in a text file. It was important to write down both the definitions and the labels as offensiveness may be stated in several ways in dictionaries. Furthermore, it was noted down whether a certain word was not included in a particular dictionary. In addition, the set of usage labels used in connection with offensive words and the definitions of these labels were collected from each dictionary. This information was presented in Table 1, Section 5.2.

Occasionally, it was slightly problematic to decide whether to include some parts of the dictionary entries or some entries as a whole in the analysis. As was stated in Section 5.4, the meanings of different words denoting the same nationality varied as these words often had several senses. The method which was applied when collecting information from dictionaries was that only those parts of the entries that mentioned a whole nationality or a large part of the nationality were included. If the entry listed several senses, only the ones with the relevant denotation were taken into account.

Sometimes there was only one sense in the whole entry that related to the particular nationality. If this denotation was only a small subgroup of a particular nationality, that entry was not taken into account, and was listed as a zero in the tables (meaning that the word in question was not included in the dictionary, at least not in the sense examined; see Section 6.2). For instance, *sawney* denotes a Scotsman (see Table 2 in Section 5.4) and not a Scottish woman. This kind of denotation was treated as comprehensive enough and it was included in the analysis. On the other hand, according to *ChD* s.v. *harp n.*, sense 3, the word *harp* denotes "an Irish American Roman Catholic". This kind of denotation was treated as too narrow, as Irish American Roman Catholics represent only a small fraction of the whole nation, and thus this sense was not taken into account in the analysis. No wider denotation signifying the Irish in general was mentioned in the entry for *harp n.* in *ChD*, and therefore *harp* in

ChD was listed as a zero in the tables. *DSUE* s.v. *harp n.*, sense 2, on the other hand, defined *harp* as "an Irish-American or an Irish person". Here, the denotation is significantly wider and therefore this part of the entry was included in the analysis.

In the dictionary analysis (Chapter 6), the information collected from the twenty dictionaries will be closely examined. Firstly, it will be seen how often and in what ways these words are categorised as offensive, and whether there are differences between single volumes or different dictionary types. The information on the frequencies of the indications of offensiveness will be presented in tabular form in Section 6.1. Moreover, it will be examined in Section 6.1 whether there are differences in the degree of offensiveness between the words for different nationalities and nationality groups: that is to say, whether the words for a certain nationality are treated as more insulting than words denoting other nationalities. It will be studied, for example, if British dictionaries label the words denoting the Irish as more offensive than the words denoting the British, as the Irish can be seen to have been oppressed by Brits in the past (see Chapter 4).

In Section 6.2, I will examine in greater detail how the offensiveness of the words has been stated. For instance, I will study whether the words have been labelled as being *often* or *sometimes* offensive or whether the pejorative marker has been modified in some other way and whether there are differences between individual words referring to the same nationality. Moreover, as opposed to Section 6.1, the treatment of individual words in dictionaries will also be examined in Section 6.2.

Some problems may arise from the fact that there are several words that are only included in a few dictionaries. Not many conclusions can be drawn from a single volume. For this reason, I have included only words that appear in at least two of the dictionaries studied. In addition, the fact that some words are not often included in dictionaries is also an interesting finding. What is more, it is somewhat problematic that older and slightly old-fashioned words (*frog*, *Kraut*, *and Mick*, for instance) are more commonly listed in dictionaries, whereas newer slang terms are not that often

included. That is why the inclusion of slang dictionaries (and also the latest editions of other dictionaries) is important, as they tend to list rarer and newer words as well.

In the corpus analysis, all ethnic slurs included in the dictionary analysis will be searched in the BNC. In all cases, both singular and plural forms of the words will be examined, as the reference may be to an individual as well as a group. The lemma which will be used is {word/N}, with the appropriate ethnic slur replacing *word*. This lemma shows both singular and plural forms of the word at the same time, and only shows words tagged as nouns. Of course, the automatically tagged system is not always correct but can be relied on in most cases. In places where it seems that the automatic tagging does not find the correct plural forms, for example, these forms will be separately checked. For instance, for some reason the lemma {rosbif/N} does not find the two tokens which can be found with the query *rosbifs*. Moreover, all spelling variants will also be examined in cases where the word can be written in multiple ways, such as *rosbif/-beef*. In all queries, case was ignored (the queries were not case-sensitive) because in the majority of cases capitalisation can occur, but is not essential. Moreover, the capitalisation of these words in the BNC and in the dictionaries varies.

It should be noted that many of the words examined are polysemous (such as *frog*), and thus sometimes many irrelevant tokens are likely to occur. The irrelevant tokens will be left out when calculating the frequencies. However, in some cases this may be somewhat problematic because the search string may produce a vast number of tokens and not all of them can be scrutinised in a thesis of this scope. The method which will be followed is that the one hundred first tokens in a randomised order will be analysed, and, based on this, a percentage will be calculated for all relevant hits. This percentage will then be applied to the number of all tokens found to find out the number of relevant instances. It goes without saying that these estimations may not be entirely accurate, but they will be accurate enough for the present purposes. The cases which are the most problematic are those where there is a word with hundreds or thousands of tokens and the nationality sense of the word is a

marginal one. This may result in zero occurrences in the sample of one hundred, even though there would be some instances among all tokens.

Furthermore, it can sometimes be problematic to decide whether an instance in the BNC is relevant or not. The context is not always helpful. For instance, the word *spi(c)k* can refer to a Spanish person but also to a Mexican. With little context, it is impossible to know whether the reference actually is to a Mexican. In addition to irrelevant tokens, there are also cases where the reference is to a particular nationality, but the word is not used as a noun. Many of the ethnophaulisms of the study can also be used as adjectives, as the following example taken from the BNC illustrates:

(9) FS1 1189 Hope you haven't acquired one of those **Yankee** accents in the States?

These instances will not be taken into account when calculating frequencies, as the dictionary analysis only focuses on the nouns.

After examining the frequencies, three words (*Boche*, *Pom* and *Yank*) will be analysed in more detail. These three words were chosen because several tokens for each word were found in the corpus, and based on the results of the dictionary analysis they represent three different levels of offensiveness. It will be examined whether a word marked as being rarely offensive, for example, is also rarely offensive in the BNC.

6. Dictionary findings

In this chapter, the dictionary findings of the study will be presented. In Section 6.1, the focus will be on frequencies: it will be examined how often the ethnic slurs of the study are classified as offensive. Different types of dictionaries will be compared with each other, and differences between single volumes will also be noted. Moreover, the frequencies of different ways of indicating the offensive nature of the words studied will be discussed. The frequencies of markers of offensiveness for different nationalities will also be examined. It will be seen whether the words denoting a particular nationality have been classified as more insulting than words referring to some other nationality. In Section 6.1, the words will be studied collectively, and the findings will be presented in tabular form. Information on the treatment of individual words in dictionaries will not be given.

In Section 6.2, the ways of expressing offensiveness will be analysed in more detail. It will be studied, for example, how often the indications of offensiveness have been modified in some way, such as stating that the word is "sometimes offensive" or "highly offensive". What is more, the differences between the treatment of different nationalities will be examined more thoroughly. As opposed to Section 6.1, the treatment of individual words in dictionaries will also be examined in Section 6.2. Again, tables will be used in order to present the findings.

6.1 Frequencies of indications of offensiveness

In this section, the frequencies of the indications of offensiveness in the dictionaries studied will be presented in tabular form. Subsection 6.1.1 will focus on the frequencies for dictionaries and in Subsection 6.1.2 the focus will be on the frequencies for different nationalities.

Some general remarks concerning the creation of the tables used in this section will be made first. Firstly, all figures in the tables have been rounded to full figures, and therefore they may sometimes add up to 99 or 101 per cent. Secondly, some abbreviations have been used in order to save

space: the abbreviation *off.* stands for the word *offensiveness*, the abbreviation *dict.* stands for *dictionaries* and *Br.* and *Am.* are abbreviations of *British* and *American*, respectively. Thirdly, retrieving information from certain entries in the dictionaries was not always straightforward. Some words, for example, were defined with the help of cross-references to other words, and no warnings were listed in those entries. In those cases, the potential warnings attached to the words used as the cross-reference were noted down. However, in some cases there was a different set of labels for the actual word studied and the word used as the cross-reference. In these cases, the labels concerning the actual word examined were taken into account instead of the warnings applied to the cross-reference entry.

6.1.1 Frequencies for dictionaries

Table 3 below is a summary of all the nationalities studied, and the focus is on dictionaries. The tables based on which Table 3 was created (tables concerning the frequencies for individual nationalities) can be found in the appendices. In Table 3 below, the number of the words found in a particular dictionary or dictionary type is listed in the column "Words included". The column "Off. not stated" presents the number of entries where there is no indication of potential offensiveness. The remaining columns include the frequencies for entries where offensiveness is mentioned in one way or another. Instances of the entries stating that the word is not always used in a pejorative manner (such as "sometimes offensive") are also included in this group, as this kind of note in a dictionary can be seen as a warning of potential offensiveness, even if the word is not insulting in every context.

Table 3 divides the words into groups based on the technique with which offensiveness is stated. In the columns "Off. stated with a label", "Off. stated in the definition" and "Off. stated with a usage note", only those cases where a single technique is used in order to communicate the offensiveness of

the words are included. For example, in the column "Off. stated with a label", the labels are the only indicator in the entry that warn the reader about the potential offensiveness of the word. If a label is combined with another technique of marking offensiveness, such as a usage note, these cases are listed in the column "Multiple indications for off.".

It should be noted that it was sometimes problematic to categorise instances under the column "Off. stated with a usage note". In this study, the term *usage note* is used to refer to the following things: firstly, if there was a sentence such as *Do not use this word* at the end of the definition, the sentence was treated as a usage note. Furthermore, if there was a word such as *offensive* after the actual definition, not italicised or otherwise marked as a separate label, it was treated as a usage note, too. If the format was, for example, "an insulting word for...", the particular entry was categorised in the "Off. stated in the definition" column (see also Section 3.2 for the different ways of indicating negative attitude in dictionaries).

Table 3 below summarises the frequencies of the indications of offensiveness used in the dictionaries studied:

Table 3. Summary: frequencies of indications of offensiveness (dictionaries)

		Words included	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these (percentage):
Br. general purpose dict.	*BED*	20	0	5	9	0	6	20 (100%)
	CED	24	10	5	9	0	0	14 (58%)
	ChD	28	11	11	6	0	0	17 (61%)
	EWEDL	17	0	3	2	0	12	17 (100%)
	ODE	24	3	21	0	0	0	21 (88%)
	OED	35	11	16	4	1	3	24 (69%)
	TOTAL	**148**	**35**	**61**	**30**	**1**	**21**	**113 (76%)**

		Words included	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these (percentage):
Am. general purpose dict.	*AHD*	15	4	1	0	0	10	11 (73%)
	EWEDNY	18	0	4	1	0	13	18 (100%)
	RHD	22	6	16	0	0	0	16 (73%)
	W3	30	13	0	0	17	0	17 (57%)
	TOTAL	**85**	**23**	**21**	**1**	**17**	**23**	**62 (73%)**
Br. learner's dict.	*CALD*	10	1	7	0	1	1	9 (90%)
	COBUILD	7	0	1	0	4	2	7 (100%)
	LDCE	13	0	0	2	3	8	13 (100%)
	MED	13	0	0	2	2	9	13 (100%)
	OALD	15	0	3	4	0	8	15 (100%)
	TOTAL	**58**	**1**	**11**	**8**	**10**	**28**	**57 (98%)**
Am. collegiate dict.	*AHCD*	14	4	0	0	0	10	10 (71%)
	MWCD	15	5	10	0	0	0	10 (67%)
	WNW	16	5	0	0	11	0	11 (69%)
	TOTAL	**45**	**14**	**10**	**0**	**11**	**10**	**31 (69%)**
Slang dict.	*DCS*	19	10	0	9	0	0	9 (47%)
	DSUE	27	14	0	9	4	0	13 (48%)
	TOTAL	**46**	**24**	**0**	**18**	**4**	**0**	**22 (48%)**
ALL DICT.	**TOTAL**	**382**	**97 (25%)**	**103 (36%)**	**57 (20%)**	**43 (15%)**	**82 (29%)**	**285 (75%)**

As can be seen from Table 3, the majority of the words examined have been marked as offensive in one way or another: 75 per cent of all entries examined include a warning of some sort, whereas 25 per cent do not. This corresponds to Norri's (2000, 91) findings: he examined twenty nationality words and twenty terms relating to race and culture in ten dictionaries, and there was an indication of offensiveness in at least sixty per cent of all the entries examined in both categories. With nationality words, the percentage of negative indications varied from 68 per cent to 94 per cent between dictionaries (ibid., p. 103). However, in the present study the degree of marking potential offensiveness varies more, from 47% in *DCS* to 100% in several of the dictionaries. It should be noted that Norri's (2000) study did not include any slang dictionaries, which are the group to include least

warnings in this study (a warning is included in 48% of the entries examined). The exclusion of slang dictionaries in Norri's (ibid.) study does not, however, account for all the differences; there are significant discrepancies between the editions of the same dictionaries in these two studies. For instance, in Norri's (ibid., p. 103) study, *W3* included a warning of some sort in 85% of the cases, whereas in this study the figure is only 57%. Of course, some differences may be explained by the fact that the words examined in these two studies were mostly different.

The dictionary group that most often marks the words examined as offensive in this study is the group of British learner's dictionaries. They indicate the potential offensiveness in 98 per cent of the cases – in fact, there is only one word that one dictionary (*CALD*) has not classified as offensive (see *CALD* s.v. *Limey n.*), otherwise there is a warning of some kind in all the other entries. Thus, all volumes in this group seem to be highly unanimous at least when examining this general summary. The high rate of warnings found in these learner's dictionaries is not surprising. In Norri's (2000, 91) study, the two learner's dictionaries included in the study were the ones to most use warnings in connection with several types of potentially insulting words. As Norri (ibid.), notes, "[i]t has probably been the compiler's intention to alert students to the risks involved in the use of some of these words".

The groups that have the second-highest rate for the indications of offensiveness in this study are British and American general purpose dictionaries. In the British general purpose dictionaries, 76% of all the entries examined are accompanied by a warning of some sort, and with the American general purpose dictionaries the percentage is 73%. It is not surprising that the rates are smaller for general purpose dictionaries than for British learner's dictionaries, as the target audience of the former consists of native speakers who may have a better sense of which words to regard as offensive than foreigners would have. However, the rates for indicating offensiveness vary from 58% and 57% to 100% with both British and American general purpose dictionaries respectively, and therefore these general purpose dictionaries are far from the unanimity of learner's dictionaries.

As concerns British general purpose dictionaries, both *BED* and *EWEDL* stand out in their high level of marking offensive words: both dictionaries include a warning of some sort in all entries examined. Both of these volumes are works by the same publisher and editor-in-chief, so their similarity is not a surprise. The policy of marking every potentially insulting nationality word as insulting can be seen as systematic, but *EWEDL* (and also *EWEDNY*) has been criticised for its way of handling potentially insulting words. According to Landau (2001, 234), "*EWED* considers almost any word offensive that has to do with mental or physical incapacity, mental mistakes, sex, age or race".

ODE is situated somewhere between the extreme ends, classifying 88% of the words as potentially offensive, whereas the *OED, CED* and *ChD* include warnings to a lesser extent. The *OED* has classified 69% of all the words as offensive, and in *CED* and *ChD* approximately 60% of all entries include a warning of some sort. This is perhaps somewhat surprising, as nationality words and ethnic slurs in general have nowadays become tabooed in the sense that they are usually treated with special care (Wachal 2002, 201). If slightly more than half of the entries are marked as offensive, the number does not reflect the development of ethnic slurs into taboos. The rate for the *OED* could be explained by the fact that some of the entries examined have not been fully updated to their third edition versions, and thus their content is older. During the course of this study, some of the *OED* entries were in fact updated to their third edition versions, and more warnings were added. This was the case with *Boche*, *Fritz*, *Heinie* and *Jerry*, which did not have warnings attached to them in their second edition entries (the *OED2* s.v. *Boche n.*; *Fritz n.1*; *Heinie n.1*; and *Jerry n.2*; see Section 6.2 for the discussion on individual words). Ethnic words were not treated with extensive discretion some decades ago (Béjoint 2010, 208).

Nevertheless, *CED* and *ChD* have low rates in expressing offensiveness, too, and therefore the publication date of a particular dictionary cannot be the only determining factor when examining the number of words marked as offensive. Furthermore, Norri's (2000, 103) study, which uses older

editions of these dictionaries, lists higher frequencies for all of these three British general purpose dictionaries when it comes to indicating the potential offensiveness of insulting nationality words. This is interesting, as one may expect to find more warnings in more recent works. However, the reason behind these differences may lie in the different selection of the words examined in these two studies. As was stated in the previous sections, the words denoting oppressed nationality groups are usually labelled as the most offensive (Landau 2001, 232; Norri 2000, 83). In the present study, many nationality groups examined represent those that have been in power: the British, the French, Germans and Americans, and this may be the reason why here the rate for the indications of offensiveness is lower in some cases than it is in Norri's (ibid.) study. The frequency rates for different nationalities will be studied in Subsection 6.1.2.

When it comes to American general purpose dictionaries, the policy of *EWEDNY* seems to be similar to its British counterpart: 100% of all entries examined have been marked as offensive. *AHD* and *RHD* both mark 73% of the words studied as offensive, whereas *W3* has an indicator of the negative connotation in 57% of the entries. This relatively low rate for *W3* could be explained by the fact that it is the oldest of the dictionaries examined, as most of its content dates from the 1960s. Nevertheless, as was stated above, the publication date clearly is not the only factor that determines the number of words marked as offensive – some newer dictionaries have an equally low rate of the indications of offensiveness. *RHD* was published more than a decade before *AHD* (see References), yet their rate for markers of insult is the same. Again, the rates for the indications of offensiveness are lower in this study than they were in Norri's (2000, 103) findings: in Norris's (ibid.) study, the rates were 94%, 90% and 85% for *AHD*, *RHD* and *W3* respectively when only nationality words are taken into account. The reasons for this are probably the same as the ones presented above when discussing British general purpose dictionaries; the words included in Norri's (ibid.) study were perhaps seen as more offensive as some of them denoted different nationalities from the ones here examined. The

frequencies of the indications of offensiveness for particular nationalities will be discussed later in this section, and this will shed some light on the matter. Moreover, the differences between individual words referring to the same nationality will be examined in Section 6.2, as there are also differences in the degree of offensiveness between words denoting the same nationality.

In collegiate dictionaries, markers of negative attitude are present in 69% of the entries studied, and the rates are rather similar in all three volumes examined. It is perhaps somewhat surprising that collegiate dictionaries use fewer warnings than some general purpose dictionaries. However, it should be noted that collegiate dictionaries are not similar to British learner's dictionaries: as was stated in Section 5.1, collegiate dictionaries are also aimed at high school and undergraduate students (see Section 5.1 for further discussion on the differences between learner's dictionaries and collegiate dictionaries), and not as much at foreign learners.

The two slang dictionaries included in the study, *DCS* and *DSUE*, have the lowest rate for the indications of offensiveness, *DCS* marking 47% and *DSUE* 48% of the entries examined as potentially insulting. Perhaps slang dictionaries are not usually seen as works that would need to provide the reader with an excessive number of usage notes and other information; slang is sometimes seen as an example of "bad language" as such, as Battistella (2005, 89) notes. Nonetheless, these two slang dictionaries do mark some of the words as offensive. The rates of less than 50% seem rather low, as all the words included in the study can be seen as potentially offensive, and are marked as such in at least some of the dictionaries studied.

In addition to the rates of marking potential offensiveness, Table 3 also sheds light on the frequencies of different ways of expressing the offensive nature of a word. Labels are the most common way of expressing potential offensive connotations, as labels have been used in more than a third of the entries studied to indicate offensiveness. However, some of the dictionaries examined do not use labels at all to warn the reader, at least not as the only technique to communicate the insulting

connotations. The dictionaries that prefer using multiple ways of expressing offensiveness (such as *AHCD*, see also Section 6.2) do use labels with some other markers of offensiveness, but some dictionaries, such as *DCS*, *DSUE*, *W3* and *WNW,* do not use labels at all. *ODE*, *MWCD* and *RHD*, on the other hand, use nothing but labels to indicate the insulting connotations of these words.

The second most common way of marking negative attitude is using multiple ways of expressing it. This was the case with 29% of the entries examined. Half of the dictionaries studied have at least in some cases used multiple indications of offensiveness, and *AHCD* even uses multiple ways of indicating offensiveness in each entry examined. Dictionaries aim at saving space and compressing their entries (Jackson 1988, 231) but this strategy makes them do the opposite. Perhaps the reason for the multiple indications of offensiveness can be explained by the importance of being politically correct and the increased concern for the rights of minorities. Section 6.2 will shed more light on the multiple ways of marking offensiveness and discuss the different combinations in detail.

In 20% of the entries, the offensiveness of the word is stated in the definition. 11 out of the 20 dictionaries have used this technique to communicate the offensive connotations of the words in at least some of their entries. *DCS* is the only dictionary to use only this techinque to warn the reader of the insulting connotations of the words.

Usage notes are the least common way of expressing insulting connotations in the dictionaries studied as they are used in 15 per cent of the entries. In fact, 12 out of the 20 dictionaries studied do not use usage notes at all. British general purpose dictionaries in particular do not seem to favour usage notes: the *OED* is the only one to use them, and, in fact, a usage note alone is used in only one of its entries studied to indicate offensiveness. *W3* and *WNW*, on the other hand, rely solely on usage notes when indicating the insulting connotations of the words examined.

As Table 3 shows, dictionaries are not usually very homogenous when it comes to marking a particular word as offensive. Only a minority of the dictionaries studied systematically use one

technique of expressing potential offensiveness. *ODE*, *MWCD* and *RHD* use labels, *W3* and *WNW* rely on usage notes and *DCS* states the offensiveness in the definitions. The majority of the dictionaries do not exploit one technique only, but use different markers in different entries. Some dictionaries, such as *CED* and *DSUE*, bounce between two different ways of expressing offensiveness, *CED* stating it either with a label or in the definition, *DSUE* mentioning it in the definition or in a usage note. The *OED*, on the other hand, uses all three of these techniques. It is understandable that different dictionaries use different techniques when indicating the offensive connotations, but one would expect to see a greater degree of consistency in single volumes.

6.1.2 Frequencies for nationalities

Table 4 below summarises the differences between the frequency rates for different nationalities (raw numbers in brackets; see Appendices for more detailed tables):

Table 4. Summary: frequencies of all indications of offensiveness in percentages (nationalities)

	The British	Americans	The Irish	The French	Germans	Mexicans
Br. general purpose dict.	72% (26)	64% (14)	81% (22)	82% (18)	77% (17)	84% (16)
Am. general purpose dict.	67% (10)	27% (3)	82% (14)	67% (6)	82% (14)	94% (15)
Br. learner's dict.	96% (22)	100% (10)	100% (7)	100% (5)	100% (5)	100% (8)
Am. collegiate dict.	63% (5)	0% (0)	100% (7)	75% (3)	71% (5)	100% (11)
Slang dict.	63% (5)	33% (2)	11% (1)	17% (1)	40% (2)	92% (11)
ALL DICT. (TOTAL)	**76% (68)**	**51% (29)**	**76% (51)**	**72% (33)**	**77% (43)**	**92% (61)**

Before analysing the findings presented in Table 4, it should be noted that because sometimes not many of the words examined were included in a particular dictionary, there was not always as much data as would have been desirable. However, the percentages were formed based on these raw numbers, as the percentages make it easier to acquire a general picture of the findings.

As can be seen from Table 4, there are differences between the rates for the indications of offensiveness when it comes to different nationalities. It was stated earlier that the mean value for all indications of offensiveness is 75%. Table 3 shows that four of the nationalities studied are quite close to the mean value: the words denoting Germans, the British, the Irish and the French, with rates of 77, 76, 76, and 72 per cent, respectively. However, the words for Americans and Mexicans stand out: only 51% of the words denoting Americans are classified as offensive, which is a low rate compared to others. Furthermore, the words for Mexicans represent the opposite end: 92% of all the words examined denoting Mexicans are marked as offensive. Thus, the rates for the indications of offensiveness for different nationalities vary from 51% to 92%.

Nevertheless, Table 4 also shows that dictionaries do not always treat one nationality unanimously. In fact, a picture of inconsistency emerges from Table 4: the dictionary types do not always seem to follow one policy, as the percentages of labelling vary to a great extent. One dictionary group, however, is quite consistent. The rates in the group of British learner's dictionaries only vary between 96% and 100%, and it was already stated in the previous subsection that there is in fact only one word in the group of learner's dictionaries that is not classified as offensive. The group which is the second most consistent is the British general purpose dictionaries: the rates vary from 64% (the rate for Americans) to 84% (Mexicans). This variation is already quite far away from the consensus of the British learner's dictionaries. Nonetheless, with American dictionaries and slang dictionaries, the variation is even greater: with American general purpose dictionaries, the rates vary from 27%

(Americans) to 94% (Mexicans). With American collegiate dictionaries, the rates vary in the greatest possible manner, from 0% (Americans) to 100% (the Irish and Mexicans). As far as slang dictionaries are concerned, the variation is almost equally striking: from 11% (for the Irish) to 92% (Mexicans).

The discrepancies between the different dictionary groups cannot at least in every case be explained with different policies for different nationalities – in theory, it would be possible for dictionaries to mark one nationality as more offensive than another. The nationality which is treated most unanimously is the group of Mexicans. The words for Mexicans have been marked as offensive very often: in British general purpose dictionaries they have been marked as insulting in 84% of all cases, and with every other dictionary type the rate is more than 90%. Perhaps the ongoing discussion on the (illegal) immigration of Mexicans into the United States makes the dictionaries encourage their users to political correctness, and racial factors may also have an effect on the high rate.

In addition to the group of Mexicans, the variation in the group for the British is not as great as it is with the other nationality groups; the rates for the British vary from 63% (American collegiate dictionaries and slang dictionaries) to 96% (British learner's dictionaries). In fact, if the British learner's dictionaries were not taken into account, the variation range would only be from 63% to 72%. Here, the frequency rates are significantly lower than the rates for Mexicans; perhaps the reason here is partly the fact that the British can be assigned the role of the oppressor in some contexts, and not the role of the oppressed.

In other nationality groups, the rates vary more. As concerns Germans, the rates vary from 40% to 100%; with the French, the range is 17%–100%; with the Irish, 11–100%, and with Americans, the rate varies from 0% to 100%. It should be taken into account, however, that slang dictionaries classify fewer words as offensive than other types of dictionaries (see Subsection 6.1.1), and this increases the variation range significantly for the Irish, and also to some extent for the French and Germans. If the slang dictionaries were not taken into account, the percentages for the Irish would be quite consistent

and the rate for the Irish would in fact be quite high. As was stated in Chapter 4, the Irish can be seen to have been oppressed in the past by the British (see Coakley 2012, 189). It is interesting that the slang dictionaries included in the study mark some words, such as the words for Mexicans, as offensive, and do not follow the same policy with the words for the other nationalities included in this study, as this may mislead dictionary users.

The treatment of the words for Americans represents well the question of the discrepancies between different dictionaries. The rate for the indications of offensiveness varies from 0% to 100%. There is a great deal of variation between dictionaries representing a similar type; the British general purpose dictionaries mark 64% of the words as offensive, whereas the American general purpose dictionaries do so with only 27% of the words. The British learner's dictionaries mark each word in this group as offensive, whereas the American collegiate dictionaries do not mark any of the words as insulting. In slang dictionaries the rate is 33%. Of course, one cannot expect to find complete unanimity between different dictionary types but certainly at least slightly more uniform policies could be predicted.

The American dictionaries classify the words denoting the British as more offensive than the words denoting Americans. This is contrary to what Rice et al. (2010, 121) argue in their study; they found out that sometimes people regard insults targeting their own group as more insulting than insults targeting some other group. However, the British dictionaries mark the words for Brits as approximately equally offensive as the words for Americans.

The frequencies presented in Table 4 can also be analysed with the help of the groups identified in the methodology chapter (see Section 5.4). The following Table 5 presents these findings:

Table 5. Frequencies of the indications of offensiveness (groups)

Group	Nationality	Words included	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these (percentage):
1	Brits	90	22	28	19	9	12	68 (76%)
	Americans	57	28	13	6	6	4	29 (51%)
	GROUP 1 TOTAL	**147**	**50 (34%)**	**41 (42%)**	**25 (26%)**	**15 (15%)**	**16 (16%)**	**97 (66 %)**
2	The Irish	67	16	16	7	6	22	51 (76%)
	GROUP 2 TOTAL	**67**	**16 (24%)**	**16 (31%)**	**7 (14%)**	**6 (12%)**	**22 (43%)**	**51 (76%)**
3	The French	46	13	13	9	5	6	33 (72%)
	Germans	56	13	13	5	7	18	43 (77%)
	Mexicans	66	5	20	11	10	20	61 (92%)
	GROUP 3 TOTAL	**168**	**31 (18%)**	**46 (34%)**	**25 (18%)**	**22 (16%)**	**44 (32%)**	**137 (82%)**

The rates for warnings vary from 66% to 82%. The lowest rate is in Group 1 (66 per cent), followed by Group 2 (76 per cent) and Group 3 (82 per cent). Based on these rates, it could be stated that words denoting the oppressors are marked as being slightly less offensive, as Group 1 includes words for the British and Americans. Nonetheless, it is the low rate for the Americans which decreases the rate, and the rate for the Brits is in fact higher or the same than the rate for some other nationalities. However, Table 5 shows that as far as Group 1 is concerned, offensiveness has been signalled with multiple indications of offensiveness quite rarely, in only 16% of the cases. With other groups, the percentages are higher. The policy of using multiple techniques to mark offensiveness can be seen as a way of emphasising offensiveness, and therefore it could be stated that sometimes slurs targeting the British and Americans are in fact seen as less offensive than slurs targeting other nationalities.

It should be noted, nonetheless, that this study consists of many words, and the words referring to one particular nationality vary: some are older, some are not very common, and each word is not found in every dictionary. As Battistella (2005, 83) notes, the degree of offensiveness of different words and expressions varies. The variation between single words naturally increases the variation between different dictionary types and nationalities, even though some general patterns can be found. In the following section, the different ways of expressing offensive connotations employed by the dictionaries studied will be analysed in more detail. The differences in the treatment of different nationalities will also be examined more thoroughly. For instance, it will be examined whether the differences between single words can account for the great amount of variation between the rates for offensiveness.

6.2 Details of the dictionary findings

This section continues the discussion on the dictionary findings, but the focus is on details and the actual ways of marking offensiveness rather than on frequencies. The findings will be presented in tabular form, and a separate table has been created to represent each nationality examined.

Some general remarks on the tables and data collection will be made first. As was the case with Table 3 in Section 6.1, the categorisation into usage notes and definitions, for example, was not entirely straightforward. The same methods which were used in Table 3 (see Section 6.1) were applied here. In addition, the words defined with the help of cross-references were treated in the same manner as in Section 6.1. The symbols used in Tables 6–11 are presented below:

Symbol	Explanation
0	The word in question is not included in the dictionary (at least not in the sense examined).
-	The word is included in the dictionary but there is no indication of potential offensiveness.

Symbol	Explanation
$+^{D}$	The word is included in the dictionary and offensiveness is stated in the definition.
$+^{L}$	The word is included in the dictionary and offensiveness is stated with a label.
$+^{U}$	The word is included in the dictionary and offensiveness is stated with a usage note.
+!	The word is included in the dictionary and offensiveness is somehow emphasised (e.g. "a very offensive word").
(+)	The word is included in the dictionary and offensiveness is indicated either in the definition, with a usage note or with a separate label (or with a combination of these). However, it is mentioned that the word can also be used in a non-offensive way (e.g. "often pejorative", "sometimes considered offensive").

It should be noted that the symbols presented above may be combined. If both a label and a usage note, for instance, are used in order to warn the reader about the offensiveness of a word, the symbol will be in the form of $+^{LU}$. Moreover, one comment should be made on the use of the symbol +!, which represents the cases where offensiveness has been emphasised. *BED*, *EWEDL* and *EWEDNY* sometimes use so-called 'quick definitions' in their entries, where the basic meaning of the word is shortly expressed in capital letters at the beginning of the entry (for example, *EWEDL* s.v. *limey n.*, sense 1, defines the word as "BRITISH PERSON"). If this quick definition reads "OFFENSIVE TERM", the symbol +! is used in the table, because using the negative connotation as the most important part of the meaning of a word can be seen as a way of emphasising the offensiveness of that particular word.

6.2.1 Findings for Group 1

This subsection discusses the details of the dictionary findings for Group 1. Words for the British will be examined first, and the findings will be presented below in Table 6. Words for the Americans will be discussed second, and the findings will be collected in Table 7, which can be found after the discussion on the words for the British.

Table 6. Words for the British

		Jock	*Limey*	*Pom*	*Pommy/-ie*	*rosbif/-beef*	*sawney*	*Taff(y)*
Br. general purpose dict.	*BED*	(+)^{L}	+!^{D}	(+)^{L}	(+)^{L}	0	0	+^{D}
	CED	-	-	(+)^{D}	(+)^{D}	-	+^{D}	-
	ChD	-	-	(+)^{L}	(+)^{L}	+^{D}	-	(+)^{L}
	EWEDL	(+)^{DL}	(+)^{DL}	(+)^{L}	(+)^{L}	0	0	+^{DL}
	ODE	(+)^{L}	-	(+)^{L}	(+)^{L}	0	0	(+)^{L}
	OED[12]	-	+^{L}	(+)^{L}	(+)^{L}	(+)^{D}	+^{D}	-
Am. general purpose dict.	*AHD*	0	-	0	+^{DL}	0	0	0
	EWEDNY	0	(+)^{DL}	(+)^{L}	(+)^{L}	0	0	0
	RHD	-	(+)^{L}	(+)^{L}	(+)^{L}	0	0	0
	W3	-	-	(+)^{U}	(+)^{U}	0	(+)^{U}	-
Br. learner's dict.	*CALD*	(+)^{U}	-	+^{L}	+^{L}	0	0	+^{L}
	COBUILD	0	(+)^{U}	(+)^{LU}	(+)^{LU}	0	0	0
	LDCE	(+)^{U}	(+)^{D}	+^{DL}	+^{DL}	0	0	+^{D}
	MED	+^{DL}	(+)^{U}	(+)^{D}	(+)^{D}	0	0	+^{DL}
	OALD	(+)^{D}	(+)^{D}	(+)^{L}	(+)^{L}	0	0	(+)^{L}
Am. collegiate dict.	*AHCD*	0	-	0	+^{DL}	0	0	0
	MWCD	0	-	(+)^{L}	(+)^{L}	0	0	0
	WNW	0	-	(+)^{U}	(+)^{U}	0	0	0
Slang dict.	*DCS*	-	(+)^{D}	(+)^{D}	(+)^{D}	0	0	0
	DSUE	-	-	(+)^{D}	(+)^{D}	0	0	-

A fact that emerges from Table 6 is that some of the words included are quite rare, or at least they are not very commonly found in the dictionaries studied. *Rosbif/-beef* and *sawney* are only found in fewer than 5 dictionaries each. Other words are more common and they are found in more than half of the dictionaries examined. Nonetheless, most of the words for the British included in this study are well represented in the dictionaries examined. The inclusion policies of the dictionaries can also partly be explained with dialectal factors: American dictionaries list fewer words denoting the British than British dictionaries. For example, *Taff(y)* and *Jock* are included in *EWEDL* (s.v. *Taffy n.* and *Jock n.*) but are not listed in *EWEDNY*. It is likely that the British use more words denoting the Scottish or the Welsh than Americans do, so these differences in the inclusion of words are understandable.

[12] The entries for *Jock*, *Limey*, *sawney* and *Taff(y)* represent the *OED2*, whereas *Pom*, *Pommy/-ie* and *rosbif/-beef* were fully updated for the *OED3*.

As far as the words that are included are concerned, the number of brackets seen in Table 6 is remarkably high. Thus, it has often been expressed that the word in question is not considered insulting in every context. For example, *COBUILD* s.v. *limey n.* defines the word as follows: "Some Americans refer to British people as limeys. Some people consider this use offensive", and *MWCD* s.v. *Pommy n.* labels the word as *usu disparaging*. In addition, there are many minus signs present in the table and thus many words are not categorised as offensive at all. Moreover, not many exclamation marks are to be seen, and thus the offensiveness of the words has not been emphasised with words such as *highly* or *very*. In fact, the only entry emphasising offensiveness in this manner is *BED* s.v. *limey n.*, sense 1, which uses the quick definition "OFFENSIVE TERM" for *Limey*.

All these factors add to the impression that the words denoting the British are not taken to be extremely offensive. Furthermore, Table 6 shows that mostly one technique has been used in order to indicate offensiveness, such as a label or a usage note. There are some cases where insulting connotations are mentioned in several ways (most often with the combination of a label and stating the negative connotation in the definition, sometimes the combination of a label and a usage note), but those cases are in the minority. This policy also further contributes to the impression that these words are not considered to be highly offensive, or at least not as offensive as words denoting some other nationality. Marking offensiveness with several techniques can be seen as a way of emphasising the offensiveness of a particular word. Of course, as was stated in Section 6.1, it can be questioned whether this kind of policy is reasonable in the first place as combining a usage note with a label and then stating the negative connotations in the definition under the same entry does not save space but the result is the opposite.

There is some variation that can be seen between the markers of offensiveness for individual words. *Jock* and *Limey* are not as often mentioned to be offensive as *Pom* and *Pommy/-ie* are, for instance. *Pom* and *Pommy/-ie* are marked as offensive in every entry, whereas *Jock* and *Limey* are

stated to be insulting in only half of them. In addition to this kind of variation between different terms, there also is variation when it comes to indicating the attitude linked to a certain word. Of course, it is only to be expected that dictionaries do not treat each word unanimously, as they are all independent works.

However, one would perhaps expect some more uniformity. *Taff(y)* can be quoted as an example. In four dictionaries there is no mention of the offensiveness of the word whatsoever. In three dictionaries, the word is stated to be offensive in some contexts. In five dictionaries, *Taff(y)* is treated as offensive in all contexts, or at least there is no mention that the connotation is only sometimes negative. However, when defining their labels for insulting words, several dictionaries state that they are used for words which can be seen as offensive or pejorative, but not necessarily (see Table 1 in Section 5.2). Thus, the definition of the label itself may imply that the word is not always used offensively. Nonetheless, it seems that many dictionaries seem to prefer to modify the label or the definition with words such as *sometimes* or *usually* even though the definition of the label states that the word is not always used offensively. For example, *BED* (p. xviii) states that the label *offensive* is used for words which are "likely to be offensive to many people, for example because of being racist or sexual". *BED* s.v. *Jock n.* labels the word as *offensive in some contexts*. The part "in some contexts" feels somewhat unnecessary, as the definition of the label itself states that words labelled as offensive are likely to be perceived as such – but not necessarily.

Even though there is some variation as to whether *Taff(y)* is marked as an insulting word or not, in those cases where the indications of offensiveness have been used they are used quite unanimously. Only the labels *offensive* and *insult* have been used, and the formulations used in the definitions are rather similar, such as "an offensive term for a Welsh person" (*BED* s.v. *Taffy n.*) or "an insulting word for someone from Wales" (*MED* s.v. *Taffy n.*). It was stated in Section 5.2 that dictionaries tend to distinguish between labels focusing on the addresser's intention and labels focusing on the reaction of

the addressee. This would suggest that *Taff(y)* is mostly taken to be offensive by the addressee, and the person saying or writing *Taff(y)* does not, at least not necessarily, mean it to be taken offensively. It can be questioned whether this in fact is the case; one would assume that a potentially offensive word is also used to hurt someone.

Some words, on the other hand, were almost systematically labelled as pejorative, not offensive. Thus, according to the dictionaries studied, some words are mostly used in order to show disapproval. *Pommy/-ie* is mostly marked to be disparaging, disapproving and derogatory, not offensive. *CED* s.v. *pommy n.* is the only dictionary to state that the word is offensive, and therefore the focus is on the reaction of the addressee in this case. Three dictionaries, on the other hand (*AHD* s.v. *pommy n.*, *AHCD* s.v. *pommy n.* and *COBUILD* s.v. *pommy n.*), mark the word to be both disparaging or disapproving and offensive. This seems reasonable, as the role of the context is essential when it comes to ethnophaulisms and insults in general; sometimes they can be used non-offensively or at least without any disparaging intent, whereas in some cases they are used to belittle a particular individual or group. This raises the question whether a word should always be labelled as both disparaging and offensive.

What is more, the way *AHD*, *AHCD* and *COBUILD* treat the word *Pommy/-ie* explains partly why dictionaries sometimes use multiple techniques to indicate offensiveness. *AHD* s.v. *pommy n.*, for instance, uses a label to indicate that the word is offensive, and states in the definition that it is also disparaging. Nevertheless, both the offensive and the disparaging connotations could be expressed with labels, or even a single label could be used in order to communicate both these negative connotations. They could also both be mentioned in the definitions and then no label would be needed. To sum up, it seems that distinguishing between offensive and pejorative words does not always seem to hold in dictionaries, as there is variation between the ways dictionaries classify ethnic slurs: whether they are offensive, disparaging or both.

Interestingly, *DSUE* s.v. *Pommy n.* mentions that “[a]lthough this word carries a definite negative connotation, it also can be used as a term of affectionate abuse”. This kind of usage is hardly ever mentioned in dictionaries – at least when examining the entries included in this study. This is perhaps somewhat surprising, as one could imagine that any one of these terms could also be used affectionately in addition to the ‘traditional’ offensive manner.

Next, the findings for the words denoting Americans will be presented. Table 7 below summarises the results:

Table 7. Words for Americans

		seppo	*septic (tank)*	*Yank*	*Yankee*	*Yanqui*
Br. general purpose dict.	*BED*	$+^{D}$	$+^{DL}$	$(+)^{L}$	$(+)^{L}$	$+^{D}$
	CED	-	0	-	$(+)^{L}$	$+^{L}$
	ChD	0	0	-	-	-
	EWEDL	0	0	$+^{D}$	$(+)^{L}$	$+^{D}$
	ODE	0	0	$(+)^{L}$	$(+)^{L}$	$(+)^{L}$
	OED[13]	0	$+^{L}$	-	-	-
Am. general purpose dict.	*AHD*	0	0	-	-	0
	EWEDNY	0	0	$(+)^{L}$	$(+)^{L}$	$+^{D}$
	RHD	0	0	-	-	-
	W3	0	0	-	-	-
Br. learner’s dict.	*CALD*	0	0	$+^{L}$	$+^{L}$	0
	COBUILD	0	0	$(+)^{U}$	$(+)^{U}$	0
	LDCE	0	0	$(+)^{U}$	$(+)^{U}$	0
	MED	0	0	$+^{DL}$	$(+)^{U}$	0
	OALD	0	0	$(+)^{DL}$	$(+)^{DL}$	0
Am. collegiate dict.	*AHCD*	0	0	-	-	0
	MWCD	0	0	-	-	-
	WNW	0	0	-	-	-
Slang dict.	*DCS*	0	-	-	-	0
	DSUE	-	$(+)^{D}$	$(+)^{U}$	0	0

Again, not all words are regularly found in the dictionaries examined. *Seppo* and *septic (tank)* are only included in a few dictionaries, *Yanqui* in approximately half of the volumes. *Yank*, on the other hand, is found in all of the dictionaries studied and there is only one dictionary (*DSUE*) which does not

[13] All entries found in the *OED* (*septic (tank)*, *Yank*, *Yankee* and *Yanqui*) were listed in the *OED2*.

include the word *Yankee*. At first, it may seem puzzling that *seppo* and *septic (tank)* are only listed in British dictionaries (and slang dictionaries) and not in American volumes even though they denote Americans. It was stated earlier in this subsection that *Taff(y)* and *Jock* (denoting a Welsh person and a Scotsman, respectively) were more commonly included in British dictionaries than in American dictionaries, and that it is understandable as the British are more likely to refer to the Welsh and the Scottish than Americans. Here, the case is not similar because *seppo* and *septic (tank)* denote all Americans, not only a subgroup and both these words represent Australian slang (*BED* s.v. *seppo n.* and *septic n.*). Americans are probably not as likely to belittle their whole nation the way some subgroup would be likely to belittle another subgroup. Moreover, Americans are not very likely to use Australian slang very often. It seems natural to find more Australian slang in British dictionaries, as many aspects of British English are still maintained in Australian and New Zealand English, and the official writing system in Australia and New Zealand is based on British English (Peters and Burridge 2012, 235–236). In fact, Peters and Burridge (ibid., p. 234) argue that Australian independence from Britain was only strongly affirmed in 1971, after Britain's final military withdrawal from Singapore.

As was the case with the words denoting the British, the words denoting Americans do not seem to be extremely offensive according to the dictionaries, either. In fact, in half of the entries examined there is no marker of potential offensiveness at all. When offensiveness is stated, many indications are modified with words such as *often* or *sometimes*. For example, *ODE* s.v. *Yank n.*, sense 1, labels the word as *often derogatory* and *EWEDNY* s.v. *Yankee n.,* sense 1, labels *Yankee* as *offensive in some contexts.* Furthermore, there are no exclamation marks in Table 7, which means that it has not been mentioned that any of these words would be highly offensive.

None of the words for Americans is treated as significantly more insulting than another. In most cases, only one technique is used in order to state the negative connotations, and labels are the most common way of expressing offensiveness. Three dictionaries (*BED*, *MED* and *OALD*) also use

combinations: in some cases, they state the offensiveness of a word in the definition and also use a separate label to mark the ethnophaulism in question as insulting. With *OALD*, this can be explained by the fact that the label is used to express that these words are often disapproving, whereas the definition states that the words are "slightly offensive" (*OALD* s.v. *Yank n.* and *Yankee n.*, sense 3). Thus, the definition and the label emphasise different things; the label focuses on the intention of the addresser, whereas the definition mentions the possible reaction of the addressee.

In *BED* and *MED*, combinations are not used because of the distinction between pejorative and offensive usage. *BED* s.v. *septic n.* gives a cross-reference to *seppo*, and labels *septic* as *offensive. Seppo* is defined as "an offensive term for a citizen of the United States" (*BED* s.v. *seppo n.*) and thus it is repeated that the word is offensive. In fact, this entry could also be categorised under the column "off. stated with a label", as the other indication of offensiveness can only be found in the cross-reference entry. However, as the cross-reference is the only definition for the word, this can be counted as a combination. What is more, *MED* has a repetitive policy, too: it labels *Yank* as *impolite* and states that it is "an insulting word for someone from the US" (*MED* s.v. *Yank n.*). As *MED* uses the label *impolite* for words that are "not taboo but will certainly offend some people" (see Table 1 in Section 5.2), the offensive connotation of the word *Yank* is mentioned twice in the same entry. This seems unnecessary as dictionaries should aim at saving space (Jackson 1988, 231).

When looking at the indications of offensiveness used for Americans, there are discrepancies when classifying individual words as offensive or pejorative. *Yanqui* is an exception as there are not that many incongruities; it is categorised as both disparaging and offensive in the definitions of *BED* (s.v. *yanqui n.*), *EWEDL* (s.v. *yanqui n.*) and *EWEDNY* (s.v. *yanqui n.*). *EWEDNY* s.v. *yanqui n.*, for example, defines the word as "an offensive term used by some members of Spanish-speaking American communities to refer disparagingly to an English-speaking U.S. citizen". *ODE* s.v. *Yanqui n.*, on the other hand, defines the word with a cross-reference to *Yankee*, and labels the word as *often*

derogatory, not as *disparaging* (*ODE* s.v. *Yankee n.*, sense 1). *CED* s.v. *yanqui n.* labels the word as *derogatory*, too.

There are more discrepancies when it comes to *Yank* and *Yankee. Yank* can be quoted as an example. When looking at the data for *Yank* in Table 7, half of the dictionaries mark the possible negative connotations of the word, and half of the dictionaries do not classify the word as offensive or disparaging. The dictionaries indicating the negative connotations are not unanimous: *EWEDNY* s.v. *Yank n.*, *BED* s.v. *Yank n.* and *COBUILD* s.v. *Yank n.* state that the word is sometimes offensive. *EWEDL* s.v. *Yank n.* and *MED* s.v. *Yank n.* argue that the word is offensive or insulting, and a modifying adverbial such as *sometimes* is not used. According to *ODE* s.v. *Yank n.*, sense 1, *DSUE* s.v. *Yank n.*, sense 1, and *LDCE* s.v. *Yank n.*, the word is often derogatory or disapproving. *CALD* s.v. *Yank n.* simply labels the word as disapproving, not including any time adverbial. According to *OALD* s.v. *Yank n.*, the word is both often disapproving and slightly offensive. Again, distinguishing between words that are offensive and words that are disparaging seems too difficult a task, and it can be questioned whether making this distinction in a dictionary is even necessary.

6.2.2 Findings for Group 2

This subsection presents the findings for Group 2, which consists of words denoting the Irish. The findings are collected in Table 8 below:

Table 8. Words for the Irish

		bogtrotter	*harp*	*Irisher*	*Mick*	*paddy*	*Pat*
Br. general purpose dict.	*BED*	$+!^{DL}$	0	0	$+!^{DL}$	$+^{D}$	$+^{D}$
	CED	$+!^{D}$	0	0	$+^{L}$	$(+)^{D}$	-
	ChD	$+^{L}$	0	$(+)^{L}$	$+^{L}$	$(+)^{D}$	-
	EWEDL	$+!^{DL}$	0	0	$+!^{DL}$	$+^{DL}$	$+^{DL}$
	ODE	$+^{L}$	0	0	$+^{L}$	$(+)^{L}$	$(+)^{L}$
	OED[14]	-	-	-	$(+)^{L}$	$(+)^{U}$	$(+)^{L}$

[14] *Bogtrotter* and *harp* were listed in the *OED2*, whereas *Irisher*, *Mick*, *paddy* and *Pat* were updated for the *OED3*.

		bogtrotter	*harp*	*Irisher*	*Mick*	*paddy*	*Pat*
Am. general purpose dict.	*AHD*	$+^{DL}$	0	0	$+^{DL}$	$(+)^{DL}$	0
	EWEDNY	$+!^{DL}$	0	0	$+!^{DL}$	$+^{DL}$	$+^{DL}$
	RHD	$+^{L}$	$+^{L}$	0	$(+)^{L}$	$(+)^{L}$	0
	W3	$(+)^{U}$	$(+)^{U}$	-	$(+)^{U}$	-	-
Br. learner's dict.	*CALD*	0	0	0	0	$+^{DL}$	0
	COBUILD	0	0	0	0	0	0
	LDCE	0	0	0	$+^{DLU}$	$+^{DLU}$	0
	MED	0	0	0	$+^{DL}$	$+^{DL}$	0
	OALD	0	0	0	$+^{DL}$	$+^{D}$	0
Am. collegiate dict.	*AHCD*	$+^{DL}$	0	0	$+^{DL}$	$+^{DL}$	0
	MWCD	0	0	0	$(+)^{L}$	$(+)^{L}$	0
	WNW	0	0	0	$(+)^{U}$	$(+)^{U}$	0
Slang dict.	*DCS*	$+^{D}$	0	0	-	-	0
	DSUE	-	-	-	-	-	-

Again, Table 8 shows that some words are only listed in a few dictionaries. *Harp, Irisher* and *Pat* are included in fewer than half of the dictionaries examined. Interestingly, *COBUILD* does not list any of the words belonging to Group 2, even though one would expect to find at least *Mick* or *paddy* in each dictionary.

When compared to Group 1, there are fewer brackets in Table 8. Thus, the words denoting the Irish are more often stated to be offensive in every case, or at least it is not mentioned that they are only sometimes insulting. Moreover, some exclamation marks are seen in Table 8 and therefore the offensiveness of these words has been emphasised in several cases. For instance, *CED* s.v. *bogtrotter n.* defines the word as "a highly offensive term for an Irish person". In addition, a great number of combinations of different techniques are used in order to express offensiveness. The most common combination is an admonitory remark in the definition and a separate label. Nonetheless, there are also other kinds of combinations: for instance, *LDCE* s.v. *Paddy n.* defines the word as "an offensive word for someone from Ireland", labels it as *taboo* and adds the note "[d]o not use this word". This kind of policy again raises the question of whether all these indications of offensiveness are necessary; a reader would probably notice the offensive connotations with fewer indications, too. The details of the

findings for the Irish show that mere percentages for the indications of offensiveness do not tell the whole truth; even though the words for the British and the words for the Irish were both marked as offensive in 76% of the entries (see Section 6.1), the words for the Irish seem to have a more prominent negative connotation according to the dictionaries.

There are some differences between the ways dictionaries treat individual words in Group 2. *Irisher* seems to be less offensive than other ethnophaulisms in this group; only *ChD* s.v. *Irisher n.* labels it as *often slightly contemptuous*, whereas the *OED3*, *W3* and *DSUE* (s.v. *Irisher n.*) do not indicate the offensiveness of the word in any way. *Harp* and *Pat* are marked as offensive in approximately half of the dictionaries including the words. *Bogtrotter*, *Mick* and *paddy* seem to be more insulting according to the dictionaries examined: an indication of offensiveness is present in the great majority of the entries.

As was the case with some words belonging to Group 1, Table 8 also shows the discrepancies when it comes to marking a particular word as offensive. *Mick* can be quoted as an example. In slang dictionaries, it is not marked as offensive at all. Five dictionaries indicate that the word is not insulting in every case. Ten dictionaries do not mention the non-offensive contexts and three of these ten dictionaries indicate that the word is very offensive. Different labels and definitions are used to state the offensiveness of the word: *Mick* is stated to be an offensive (e.g. *ChD* s.v. *mick n.*, sense 1), insulting (e.g. *MED* s.v. *mick n.*), disparaging (e.g. *AHD* s.v. *mick n.*), derogatory (e.g. *WNW* s.v. *mick n.*) and taboo word (e.g. *OALD* s.v. *Mick n.*). Again, different dictionaries are far from being unanimous, but the variation between different works is to some extent to be expected.

Interestingly, *DCS* s.v. *Paddy n.* mentions that "[a]n alternative, usually with a slightly more pejorative emphasis, is mick". Nevertheless, in its entry for *Mick* (*DCS* s.v. *mick n.,* sense 1a) there is no mention of the insulting connotations of the word whatsoever. This shows that in addition to the discrepancies between different volumes, there are also incongruities within a single dictionary.

6.2.3 Findings for Group 3

This subsection presents the findings for Group 3. Words for the French will be examined first, and the findings will be presented below in Table 9. Words for Germans will be discussed second, and the findings will be collected in Table 10. Words for Mexicans will be examined after the words for Germans, and the findings will be summarised in Table 11.

Table 9. Words for the French

		Frencher	*Frenchy/-ie*	*frog*	*frog-eater*	*froggy/-ie*	*mounseer*	*parleyvoo*
Br. general purpose dict.	*BED*	0	0	$+^{D}$	0	0	0	0
	CED	0	-	$+^{D}$	0	$+^{D}$	0	-
	ChD	0	$+^{D}$	$+^{D}$	$+^{D}$	$+^{D}$	-	-
	EWEDL	0	0	$+^{DL}$	0	0	0	0
	ODE	0	$(+)^{L}$	$+^{L}$	0	$+^{L}$	0	0
	OED[15]	$+^{L}$	$+^{L}$	$(+)^{L}$	$+^{L}$	$(+)^{L}$	$(+)^{L}$	$(+)^{L}$
Am. general purpose dict.	*AHD*	0	0	$+^{DL}$	0	0	0	0
	EWEDNY	0	0	$+^{DL}$	0	0	0	0
	RHD	0	-	$+^{L}$	0	0	0	0
	W3	$(+)^{U}$	-	$(+)^{U}$	$(+)^{U}$	0	0	-
Br. learner's dict.	*CALD*	0	0	$+^{L}$	0	0	0	0
	COBUILD	0	0	$(+)^{U}$	0	0	0	0
	LDCE	0	0	$+!^{DLU}$	0	0	0	0
	MED	0	0	$+^{DL}$	0	0	0	0
	OALD	0	0	$+^{D}$	0	0	0	0
Am. collegiate dict.	*AHCD*	0	0	$+^{DL}$	0	0	0	0
	MWCD	0	0	$(+)^{L}$	0	0	0	0
	WNW	0	-	$+^{U}$	0	0	0	0
Slang dict.	*DCS*	0	0	-	0	-	0	0
	DSUE	0	-	-	$+!^{D}$	-	0	0

Table 9 shows that many of the words are quite rarely mentioned in the dictionaries studied. *Frog* is included in every dictionary, but all other words are listed in fewer than half of the dictionaries examined. *Frencher* and *mounseer*, for example, are only included in two dictionaries, and for this reason generalisations are difficult to make for these words. Many words, such as *froggy/-ie* and *mounseer* are more commonly listed in British than in American volumes. This is understandable as

[15] All words for the French were updated for the *OED3*.

the British share a great amount of history with the French (see Chapter 4), whereas Americans and the French have not been in an equally close contact.

Overall, the degree of offensiveness in this group is rather similar to the words denoting the Irish; the words denoting the French are not always classified as offensive, but in the majority of cases they are. In addition, the indications of offensiveness are not as often modified with words such as *usually* or *sometimes* as was the case with the words denoting the British and Americans. However, when compared to the Irish, the offensiveness of the words for the French is not as often emphasised with words such as *highly*; here, this method has only been used in two entries.

Combinations of different techniques of indicating the negative connotations are used occasionally, but not as often as in Group 2. The most common combination is a remark in the definition and a separate label; in fact, the only combination not following this pattern is found in *LDCE* s.v. *frog n.*, sense 3, which uses the combination of all three techniques: a usage note, a label and stating offensiveness in the definition. The combinations seem unnecessary in several places, as the different techniques mostly repeat the same information. *EWEDL* s.v. *frog n.*, for instance, defines the word as "an offensive term for a French person" and labels the word as *offensive*. This kind of policy seems redundant. In some dictionaries, on the other hand, the use of combinations seems more justified: *AHD* s.v. *frog n.*, sense 9, and *AHCD* s.v. *frog n.*, sense 9, use the label to state that the word is offensive and mention in the definition that the word is also disparaging.

Even though some of the words in Table 9 are only found in a few dictionaries and thus generalisations concerning these words should be taken with a certain amount of reservation, it can be stated that some words in this group are marked to be more insulting that others. The words including the root[16] *frog* (and also the word *Frencher*) appear to be more offensive than *Frenchy/-ie*, *mounseer* and *parleyvoo*. For instance, only the *OED3* s.v. *parleyvoo n.* and *adj.,* sense A2, labels the word as

[16] A root is a single morpheme which is not analysable into smaller parts and which is not an affix (Herbst 2010, 86).

frequently derogatory. The three other dictionaries including *parleyvoo* do not mention the insulting connotation at all. *Frog-eater*, on the other hand, is stated to be offensive or disparaging by all four dictionaries including the word.

As was the case with *Mick* in Group 2, here the treatment of *frog* best demonstrates the discrepancies between different dictionaries. As Table 9 shows, the word has sometimes been classified as offensive, sometimes not, in some entries the offensiveness of *frog* has been emphasised and some dictionaries state that the word is not always offensive. The word has been described to be offensive (e.g. *EWEDNY* s.v. *Frog n.*), insulting (*MED* s.v. *frog n.*, sense 2), disparaging (e.g. *RHD* s.v. *frog n.1*, sense 4), derogatory (e.g. *ODE* s.v. *frog n.1*, sense 2), contemptuous (*ChD* s.v. *frog n.1*, sense 3) and taboo (*LDCE* s.v. *frog n.*, sense 3). Of course, it is natural that the exact wordings are not similar in all dictionaries, but there is again the problem that the dictionaries are not unanimous as to whether the word is offensive, pejorative or both.

Next, words denoting Germans will be discussed with the help of Table 10:

Table 10. Words for Germans

		Boche	*Fritz*	*Heinie*	*Jerry*	*Kraut*	*sausage*
Br. general purpose dict.	*BED*	$+^{D}$	0	0	$+^{DL}$	$+^{D}$	0
	CED	$+^{L}$	0	0	-	$+^{D}$	0
	ChD	$+^{L}$	0	0	-	$+^{L}$	0
	EWEDL	$+!^{DL}$	0	0	$+^{DL}$	$+^{DL}$	0
	ODE	-	-	0	$+^{L}$	$+^{L}$	0
	OED[17]	$(+)^{L}$	$(+)^{LU}$	$(+)^{LU}$	$(+)^{LU}$	$(+)^{L}$	-
Am. general purpose dict.	*AHD*	$+^{L}$	0	0	-	$+^{DL}$	0
	EWEDNY	$+!^{DL}$	0	0	$+^{DL}$	$+^{DL}$	0
	RHD	$+^{L}$	$(+)^{L}$	$+^{L}$	-	$+^{L}$	0
	W3	$(+)^{U}$	$(+)^{U}$	$(+)^{U}$	-	$(+)^{U}$	$(+)^{U}$
Br. learner's dict.	*CALD*	0	0	0	0	0	0
	COBUILD	0	0	0	0	0	0
	LDCE	0	0	0	0	$+!^{DLU}$	0
	MED	0	0	0	$+^{DL}$	$+^{DL}$	0
	OALD	0	0	0	$+^{DL}$	$+^{DL}$	0

[17] *Sausage* represents the *OED2* and all the other words were updated for the *OED3*.

		Boche	*Fritz*	*Heinie*	*Jerry*	*Kraut*	*sausage*
Am. collegiate dict.	*AHCD*	$+^{DL}$	0	0	-	$+^{DL}$	0
	MWCD	0	0	0	-	$(+)^{L}$	0
	WNW	$+^{U}$	0	0	0	$+^{U}$	0
Slang dict.	*DCS*	0	0	0	0	-	0
	DSUE	0	-	-	$+^{D}$	$(+)^{D}$	0

The words which are not included in many dictionaries are *sausage* (listed in two dictionaries), *Heinie* (in four dictionaries) and *Fritz* (in five dictionaries). *Boche*, *Jerry* and *Kraut* are more common. Here, there are no clear differences in the inclusion of words between British and American dictionaries.

The words denoting Germans seem rather similar to the words denoting the Irish in terms of offensiveness: the words are not always classified as offensive, but in the majority of the cases they are. Moreover, the indications of offensiveness are not very often modified with words such as *usually* or *sometimes*, as the rather low number of brackets in Table 10 shows. Offensiveness is emphasised with words such as *highly* in two entries. Combinations of the indications of offensiveness are used quite frequently, and again the most common combination is a label and a statement concerning the offensiveness of a word in the definition. Again, there is unnecessary repetition in some of these combinations: *EWEDNY* s.v. *Kraut n.*, for instance, labels the word as offensive and defines it as "an offensive term referring to a German". In some cases, the label and the definition provide information on different negative connotations: *AHD* and *AHCD* continue their policy of describing the word as offensive with a label and stating that it is disparaging in the definition (see *AHD* and *AHCD* s.v. *kraut n.*, sense 2). *OALD* s.v. *Kraut n.*, and *LDCE* s.v. *kraut n.*, on the other hand, use a label to state that *Kraut* is a taboo word and state in the definition that it is also offensive. Interestingly, Landau (2001, 232) argues that an old-fashioned ethnic slur "is no longer seen as much of a threat" and gives *Kraut* as an example among others. Nevertheless, based on the findings presented in Table 10, *Kraut* is still seen as a threat to some extent, as it is marked as an offensive or a disparaging term in all entries but one, and the indications of offensiveness are rarely modified with words such as *sometimes.*

The words which are marked to be the most offensive in the dictionaries examined are *Kraut* and *Boche*, as the entries not classifying these ethnic slurs as insulting are in the clear minority. *Jerry*, on the other hand, is marked as offensive in approximately half of the entries examined. Therefore, this again shows that there are differences between the offensiveness of individual ethnic slurs even when they denote the same nationality. This is also noted in the *OED3*: in its entries for *Jerry* (the *OED3* s.v. *Jerry n.3* and *adj.3*, sense A2) and *Fritz* (the *OED3* s.v. *Fritz n.1* and *adj.*, sense A1), it is stated with usage notes that these words are less derogatory than *Kraut* and *Boche*.

Again, Table 10 illustrates many discrepancies between dictionaries when it comes to individual words. *Boche*, *Kraut* and *Jerry* all represent this. The words are sometimes marked to be offensive, sometimes not. *Boche* and *Kraut* are occasionally mentioned to be very offensive, but in some entries it is stated that they are not offensive in each context. Moreover, the dictionaries examined are not unanimous when it comes to marking a particular word as offensive or pejorative; *Boche*, for example, has been classified as offensive by e.g. *AHD* s.v. *Boche n.*, as derogatory by *CED* s.v. *Boche n.*, for instance, and as both offensive and disparaging by *AHCD* s.v. *Boche n.* Thus, even though *AHD* and *AHCD* belong to the same dictionary series, there are discrepancies between these two volumes.

Next, the words denoting Mexicans will be examined with the help of Table 11:

Table 11. Words for Mexicans

		beaner	*greaseball*	*greaser*	*spic/spi(c)k*	*taco bender*	*wetback*
Br. general purpose dict.	*BED*	0	$+!^{DL}$	0	0	0	$+!^{DL}$
	CED	0	0	0	$+^{D}$	0	$+^{L}$
	ChD	0	$+^{L}$	$+^{L}$	$+^{L}$	0	-
	EWEDL	0	0	0	0	0	$+!^{DL}$
	ODE	$+^{L}$	$+^{L}$	$+^{L}$	$+^{L}$	0	$+^{L}$
	OED[18]	$+^{L}$	$+^{D}$	-	$+^{D}$	0	-
Am. general purpose dict.	*AHD*	0	$+^{DL}$	$+^{DL}$	$+^{DL}$	0	$+^{DL}$
	EWEDNY	0	$+!^{DL}$	$+!^{DL}$	$+!^{DL}$	0	$+!^{DL}$
	RHD	0	$+^{L}$	$+^{L}$	$+^{L}$	0	$+^{L}$
	W3	0	$(+)^{U}$	$(+)^{U}$	$(+)^{U}$	0	-

[18] The word *beaner* is a new entry in the *OED3*, whereas *greaseball*, *greaser*, *spic/spi(c)k* and *wetback* were listed in the *OED2* already.

		beaner	*greaseball*	*greaser*	*spic/spi(c)k*	*taco bender*	*wetback*
Br. learner's dict.	*CALD*	0	0	0	+^{L}	0	0
	COBUILD	0	0	0	0	0	+^{L}
	LDCE	0	0	0	+!^{DLU}	0	+!^{DLU}
	MED	0	0	0	0	0	+^{DL}
	OALD	0	+!^{DL}	0	+!^{DL}	0	+^{DL}
Am. collegiate dict.	*AHCD*	0	0	+^{DL}	+^{DL}	0	+^{DL}
	MWCD	0	(+)^{L}	(+)^{L}	(+)^{L}	0	(+)^{L}
	WNW	0	+^{U}	+^{U}	+^{U}	0	+^{U}
Slang dict.	*DCS*	+!^{D}	+^{D}	+!^{D}	+!^{D}	+^{D}	-
	DSUE	+^{D}	+^{D}	+^{U}	+^{U}	+^{U}	+^{D}

The rarest words to be found in the dictionaries among the words denoting Mexicans are *taco bender* (included in two dictionaries) and *beaner* (listed in four dictionaries). Other words are listed in more than half of the dictionaries studied. Some regional differences are again to be seen: *EWEDNY* includes *greaser*, *greaseball* and *spic/spi(c)k*, whereas *EWEDL* does not. *EWEDL* s.v. *spic n.* defines the word as "a highly offensive term referring to a Spanish or Italian person", and no Mexicans are mentioned. This is natural as Americans are more in contact with Mexicans than the British, and in a wider sense *spic/spi(c)k* can refer to a person "of Latin origin, (an) Italian or Hispanic" (*DCS* s.v. *spick*, *spic n.* and *adj.*).

As was stated in Section 6.1, the words for Mexicans are marked as offensive more often than words denoting the other nationalities in the present study. Often, different techniques are used in the same entry to indicate offensiveness (usually the combination of a label and a statement concerning the negative connotation in the definition), and occasionally offensiveness is emphasised. According to *OALD* s.v. *greaseball n.*, for example, *greaseball* is "a very offensive word for a person from southern Europe or Latin America". Furthermore, it is implied in many entries that the word is offensive in every context; not many brackets are seen in Table 11. As a matter of fact, the only two dictionaries modifying their warnings with *usu* (usually) are *W3* and *MWCD*, which in fact always use the abbreviation *usu* in connection with the ethnophaulisms included in the study; it is never left out when

offensiveness is mentioned (see Tables 6–11). This raises the question whether other dictionaries should follow the same policy, as offensiveness is always context-dependent. Of course, it should be taken into account that in their definitions of labels for offensive words, some dictionaries do state that the words labelled as such are likely or usually taken to be offensive, and thus not necessarily insulting in all contexts. However, it was already stated in Subsection 6.2.1 that despite their definitions of labels (stating that a word classified as offensive may not be offensive in every context), dictionaries may still add a modifier such as "in some contexts" to the label when they use these labels in dictionary entries, and thus repeat the information that has already been mentioned in the definitions of labels in the front matter of the dictionary.

Even though most ethnophaulisms targeting Mexicans are classified as offensive, the words for Mexicans are not treated completely unanimously, either. Nevertheless, the unanimity is greater within this nationality group; almost every entry has an indication of offensiveness. There are, of course, differences between the ways the offensiveness of these ethnic slurs is stated and also in the degree: according to some dictionaries, not all these words are always offensive, and some dictionaries state certain words to be extremely offensive.

It is interesting that *wetback* is not marked as offensive as often as the other words denoting Mexicans. Of course, it has been argued several times earlier in this chapter that dictionaries do not treat the same word unanimously. However, as an indication of offensiveness is present in all other entries for Mexicans (except for the *OED2* s.v. *greaser n.*, sense 2a), one would assume the treatment of *wetback* to be more unanimous. Perhaps one reason for the incongruities can be found in the definition of *wetback*: according to *ChD* s.v. *wetback n.*, sense 1, a *wetback* is "someone illegally entering the USA from Mexico by wading or swimming the Rio Grande". If somebody is claimed to have entered the United States illegally, perhaps the compilers of the dictionary assume that the illegalness implies that this word is used pejoratively and can cause offence.

The policy of slang dictionaries is interesting when it comes to words denoting Mexicans. With other nationalities, offensiveness is not often stated in slang dictionaries (except for the British, where the majority of the words are stated to be usually pejorative) but with Mexicans the words are stated to be insulting in all entries but one. This policy seems inconsistent. Mexicans seem to be one of the nationality groups that are treated with special care in dictionaries. As was stated in Section 6.2, perhaps the ongoing discussion on the (illegal) immigration of Mexicans into the United States makes the dictionaries encourage their users to political correctness. Racial factors may also have an effect on the high rate – all the other nationalities included in the study are typically seen as predominantly white.

7. Corpus findings

The detailed dictionary analysis in Section 6.2 showed that many words included in the present study are listed in only a few dictionaries. A corpus analysis using the BNC (see Section 5.3) was executed in order to find out whether the corpus data corresponds to the inclusion policies of the dictionaries: that is to say, whether the words not commonly listed in the dictionaries studied have few tokens in the corpus, too. The frequency information will be presented in Section 7.1. In addition, three ethnophaulisms (*Boche*, *Pom* and *Yank*) will be examined more thoroughly. It will be seen whether the indications of offensiveness in dictionaries for these ethnic slurs correspond to corpus information. For instance, *Boche* has been categorised as an insulting word in almost all the dictionaries examined, and it will be investigated whether the corpus data supports this information. These findings will be presented in Section 7.2.

7.1 Frequencies of the nationality words in the BNC

This section focuses on the frequencies of the ethnic slurs in the BNC. The findings are presented in the following table (Table 12). The numbers marked with an asterisk are figures that have been formed based on estimations (see Section 5.5), and these estimations have been rounded to full figures. The calculations based on which these estimations have been constructed can be found in Appendices (Appendix 2).

Table 12. Frequencies in the BNC

Group	Nationality	Word	Tokens	The number of dictionaries in which the word is included (out of 20)
Group 1	The British	*Jock*	*12	14
		Limey	8	20
		Pom	16	18
		Pommy/-ie	1	20
		rosbif/-beef	2	3
		sawney	0	4
		Taff(y)	1	12
	Americans	*seppo*	0	3
		septic (tank)	0	4
		Yank	*95	20
		Yankee	*27	19
		Yanqui	3[19]	11
Group 2	The Irish	*bogtrotter*	1	13
		harp	*0	4
		Irisher	0	4
		Mick	*15	18
		paddy	*0	19
		Pat	*0	9
Group 3	The French	*Frencher*	0	2
		Frenchy/-ie	2	8
		frog	*47	20
		frog-eater	0	4
		froggy/-ie	0	6
		mounseer	0	2
		parleyvoo	0	4
	Germans	*Boche*	9	12
		Fritz	*2	5
		Heinie	0	4
		Jerry	*16	15
		Kraut	12	18
		sausage	*10	2
	Mexicans	*beaner*	1	4
		greaseball	0	13
		greaser	1	12
		spic/spi(c)k	3	16
		taco bender	0	2
		wetback	1	19

[19] All three tokens come from the same part of the same text.

As Table 12 shows, many words included in the study were not found in the BNC at all. These words are *sawney* (for the British); *seppo* and *septic (tank)* (for Americans); *harp*, *Irisher*, *paddy* and *Pat* (for the Irish); *Frencher*, *frog-eater*, *froggy/-ie*, *mounseer* and *parleyvoo* (for the French); *Heinie* (for Germans) and both *greaseball* and *taco bender* for Mexicans. The fact that some of these words were not found in the BNC is not surprising; for instance, *taco bender* and *mounseer* were only listed in two dictionaries, *seppo* in three (see Section 6.2). However, it is surprising to see *paddy* among these words: it was included in all the dictionaries studied except for *COBUILD*. Nonetheless, the explanation for this may be that not all 899 tokens for *paddy* were examined individually, only one hundred. *Paddy* can also denote rice paddies (the *OED3* s.v. *paddy n.1*), and can be used as a name or a nickname for *Patrick*, for example, and these senses were more common in the BNC. In fact, if the noun is searched only in its plural form (as this leaves out most tokens for names), some instances are found where the reference is to the Irish, such as the following example (10):

> (10) HTG 2643 The thing about the **Paddies** is that some of them can pass as English if they've been brought up here.

Calculations may also be the reason for the low frequency of *Taff(y).*

Nevertheless, in some cases the low number of tokens cannot be explained by calculations, as estimations have not been applied. For instance, there is only one token for *Pommy/-ie*, even though the word is included in all the dictionaries studied. The reason for this may partly be that *Pommy/-ie* represents Australian and New Zealand English (*BED* s.v. *pommy n.*), and the BNC is a corpus of British English. However, the dialectal aspect cannot be the only factor as there are 16 tokens for *Pom*, which also represents Australian and New Zealand English (*BED* s.v. *pom n.*). On the other hand, there were several instances where *Pommy/-ie* was used as an adjective, such as the following:

> (11) ATE 398 Nothing could have quicker distanced him from the **Pommie** chairman or won him their attention so soon.

Dialectal reasons can at least partly explain why there is only one token for *wetback*, even though it is included in nineteen of the dictionaries studied. *Wetback* belongs to American English (*OALD* s.v. *wetback n.*). It should also be kept in mind that some words for Mexicans were more common in American than in British dictionaries. Nonetheless, the representation of words for Mexicans is rather low in the BNC in general. Perhaps Mexicans simply are not as often talked about in British contexts as the other nationalities studied, because the British do not interact with Mexicans that often due to the geographical distance, for instance (see Chapter 4).

Moreover, dialectal factors cannot be the reason for the low number of tokens for *bogtrotter*, because the word has been labelled as belonging to British English in at least some of the dictionaries studied (see *DCS* s.v. *bog-trotter n.*, for instance). The word is not classified as dated in any of the dictionaries studied, and thus it is difficult to explain why the corpus data does not correspond to dictionary findings (the word was also searched in the BNC in the form *bog-trotter*, which yielded the only token found).

When looking at the words occurring several times in the BNC, most of these are also words listed in (almost) all dictionaries, such as *frog*, *Yank*, *Yankee*, *Mick* and *Pom*. However, there are some words that have more occurrences that one would have expected based on the dictionary findings. *Sausage* was listed in two dictionaries only and yet it is estimated to have ten tokens in the BNC. However, this is an estimation and it may be that the one token found in the group of one hundred random tokens is the only occurrence in the BNC.

To sum up, the inclusion of words in dictionaries corresponds quite well to the corpus findings. Calculated estimations may have an effect on the mismatch between the dictionary and corpus data, and the mismatch may also be created by dialectal factors or the rare contact between Britain and the nationality referred to. With *bogtrotter*, however, no explanation for the mismatch was found.

7.2 *Boche*, *Pom* and *Yank* in the BNC

In this section, *Boche*, *Pom* and *Yank* in the BNC will be examined in more detail to see whether the dictionary findings correspond to the corpus data. Subsection 7.2.1 will focus on *Boche*, Subsection 7.2.2 will shed light on the tokens for *Pom*, and *Yank* will be examined in Subsection 7.2.3.

7.2.1 *Boche* in the BNC

The word *Boche* was stated to be offensive in the majority of the entries in the dictionaries studied, and the indications of offensiveness were hardly ever modified with words such as *sometimes* or *usually* (see Table 10 in Section 6.2). In two entries, the offensiveness of the word was emphasised. Thus, one would expect that the word would be used offensively in most examples found in the BNC. Nevertheless, this is not exactly the case. *Boche* has nine tokens in the BNC. Two occurrences are explicitly offensive, quoted in examples (12) and (13) below:

> (12) AMC 794 I didn't care two pins about killing **Boche**, but I did get upset at the spectacle of starving French civilians and Arabs when we first took Tunis.
> (13) HTG 321 '**Boche**, Hun, bloody German.

However, in other occurrences the use of *Boche* was not as clearly insulting as in (12) and (13) quoted above. Reference was often to an enemy in the war, but the offensiveness of the word was not stated explicitly, as (14) and (15) below show:

> (14) A61 1041 All this damage can be repaired when we have beaten the **Boche**.
> (15) HTW 3368 My master, the Comte de Beaumont, went to England to fight the **Boche**.

(14) and (15) are more neutral than the first two examples. Of course, it should be noted that an ethnic slur can be treated as an insult as such, and the sentence surrounding it does not always emphasise the offensiveness of the word. In addition, as Landau (2001, 233–234) notes, the target of the insult does not always show that s/he has been offended. Moreover, there is sometimes no reaction available in the

corpus context, as is the case if reference is to the nationality in general, not to someone who is present in the situation. This is the case in both (14) and (15). Nevertheless, in the following example (16) it seems that *Boche* is used with a more positive than a negative tone:

> (16) HWA 2272 Try to get to Paris before the blessed **Boche**.

(16) can be seen to illustrate the fact that ethnic slurs are not always used offensively and pejoratively. Therefore, the dictionary information does not completely correspond to the corpus data. However, it should be taken into account that there is not a sufficient amount of corpus context to examine whether (16) is in fact a purely positive instance, and not irony, for example. Nonetheless, even though *Boche* was marked to be offensive practically always in the dictionaries, the corpus data shows that a modifier such as *usually* could have been used more often. Even though the corpus sample was small, there were several instances where the word seemed to have a more neutral connotation, and even one instance with a possible positive connotation. This again raises the question of whether dictionaries should always apply a word such as *usually* or *sometimes* in connection with ethnic slurs.

7.2.2 *Pom* in the BNC

In the dictionaries studied, *Pom* was in most cases treated as a usually or an often offensive word, and thus one would expect to find instances in the BNC where the word has not been used offensively, but also many cases where there is a negative connotation present. This is indeed the case. There are several instances where the connotation of the word seems to be clearly negative, such as the two examples below:

> (17) CRY 1531 In person and style, he remained the down-to-earth Australian, caustic about the class-ridden antique snobberies of the '**poms**'.
> (18) J1E 2005 Foreigners tend to see him as a 'whingeing **pom**, brit etc.' and do not like the program.

In (17) and (18), *Pom* is used offensively. However, there are more neutral examples in the BNC, too, such as the following:

(19) ECU 1327 From 1974 until 1988, the **Poms** didn't win a Test Match.
(20) ECU 1335 I HAVE TWICE SEEN the Australians try to beat the **Poms** at Wembley.

(19) and (20) are more neutral than (17) and (18), even though it should be noted that, as stated previously, the word can be an insult even though the sentence itself is neutral, as there is not always enough context available when corpora are analysed. A closer look at the context of (19) and (20) reveals that the topic is rugby in both cases. Perhaps a further study on sports language could be executed, in order to see whether ethnophaulisms can be used in a more neutral way when it comes to sports, or rugby in particular.

Nevertheless, one instance of *Pom* with a clearly positive connotation can be found in the BNC, and this supports the information found in dictionaries; *Pom* is often an offensive word, but not always.

(21) EB3 827 'The **Poms** are good sorts,' he replied.

(21) illustrates that *Pom* is indeed not always used pejoratively. Of course, the addressee could still feel hurt, but the message is meant to be positive. With *Pom*, the dictionary information was similar to the corpus findings.

7.2.3 *Yank* in the BNC

Yank was classified as offensive in only half of the dictionaries studied (see Section 6.2). In the entries where offensiveness was stated, it was often indicated that the word is not always insulting. Thus, if the data of the BNC is similar to the dictionary data, there should be plenty of examples where *Yank* is used in a neutral or even positive context. Several occurrences in negative contexts should also be found.

The following instances are examples of the negative and disparaging connotations sometimes connected to *Yank:*

> (22) F9C 2708 Bloody **Yanks**.'
> (23) B3F 458 But then to us they all had the same accent — they were just **Yanks**, chewing gum, blocking our paths wherever we went, aggravatingly sure of themselves and their instant appeal to all females who hadn't actually got one foot in the grave.

(22) and (23) show that clearly the warnings found in some of the dictionaries have been accurate; *Yank* can indeed be used offensively. In fact, the message seems to be offensive in quite a few instances, which raises the question of whether more dictionaries should have warned the readers about the use of the word. However, offensiveness is not present in all tokens found in the BNC. The following instances seem rather neutral:

> (24) FS0 995 Brian and I recounted the awful journey, but, of course, all the **Yanks** had been through the same ordeal.
> (25) KCS 790 if the **Yanks** of left it a bit later they'd of missed all the action

(24) and (25) are more neutral than (22) and (23) mentioned earlier. Again, there is also at least one clearly positive use of the word *Yank*, illustrated by the following example:

> (26) HWM 163 'I quite like the **Yanks**.'

(26) demonstrates how this word, as was the case with *Boche* and *Pom*, too, can be used in a positive context. To sum up, the dictionary entries for *Yank* seem quite appropriate when compared to the corpus findings, even though there were quite a few instances in the BNC where the connotation was clearly negative.

8. Discussion

Overall, 75% of all the ethnic slurs examined were marked as offensive. There was variation between different dictionary types, single volumes within the same dictionary type, between the indications attached to words denoting different nationalities and also between single words referring to the same nationality. Variation was to be expected, but occasionally it was even greater than was anticipated. However, as Norri (2000, 92–93) points out, discrepancies are always to be expected when works by different lexicographers and publishers are compared to each other. Landau (2002, 233) notes that "there are no agreed-upon criteria for finding some usages offensive or contemptuous or abusive", and this explains why different dictionaries may not treat the same ethnophaulism in a similar manner. However, one would expect to find consistency within single volumes (Norri 2000, 93), and occasionally this consistency was lost.

The dictionary type which treated ethnic slurs most uniformly was British learner's dictionaries, attaching a warning of some sort to almost all entries examined. This was not a surprise; Norri (2000, 91) found out in his study that learner's dictionaries tend to use warnings more often than other dictionaries. Nonetheless, in the other dictionary groups some volumes also marked each of the words examined to be offensive, and in some places it was questioned whether all these warnings were necessary - especially if the same information was repeated multiple times in a single entry. As Landau (2001, 234) argues, dictionaries may nowadays even overuse the indications of offensiveness.

The most common way of indicating negative connotations was the use of labels, followed by the use of combinations of different markers, then stating the offensiveness of a word in the definition, and, finally, mentioning the insulting connotations of a word in a usage note. It was slightly problematic that many dictionaries did not give any definitions for the labels they used. This problem is also noted by Norri (2000, 93). Not defining the labels used causes problems because each dictionary uses labels in their own way, and the same label can denote different things in different

dictionaries. Atkins and Rundell (2008, 231) argue that there is "quite a lot of work involved in putting together a consistent policy on labels in a dictionary". This indeed seems to be the case.

The second most common way of indicating offensiveness was the use of several warnings in the same entry. Even though sometimes the use of these different techniques was motivated by the distinction into offensive and pejorative uses, in several places the combinations only resulted in unnecessary repetition. Dictionaries should aim at saving space where possible (Jackson 1988, 231), and repeating the same information several times in a single entry is not of much help. In addition, there was also repetition when the dictionaries defined the labels they use, and later repeated the information in their entries. For instance, many of the definitions of labels used with offensive words stated that they are attached to words that are often used to cause offence, but the words may also be used non-offensively. Nevertheless, the dictionaries modified their definitions of ethnophaulisms with expressions such as "an often offensive word" or "a usually insulting term" and thus ended up repeating the information that was already stated. However, it may well be the case that not every dictionary user reads through the front matter of the dictionary and this is the reason why the same information is repeated.

The majority of the dictionaries did not use only one technique to express offensiveness; sometimes they stated it in the definition, and sometimes they used a label, for instance. Many dictionaries bounced between different techniques, even though it would be easier for the dictionary user if s/he knew where to look for markers of negative attitude. Thus, this is an inconsistency within a dictionary, something that should not be there. In addition, the dictionaries were far from being unanimous when dividing words into offensive and disparaging ones, and those that could be seen as both. This was also noted by Norri (2000, 92): although offensive and disparaging uses are often distinguished from each other in the front matter of the dictionary, "the nuance is difficult to maintain in practice". Jackson (1988, 198–199) notes that lexicographers often include some unnecessary

information in their entries. Perhaps the general public could do without any specialisations of offensiveness; it might be enough for them to know that the word in question has to be used with caution, not whether the word is offensive or disparaging, for instance. Furthermore, some dictionaries always modified their markers of offensiveness with words such as *usually*. It can be asked if this should be the case in each dictionary: after all, as Burridge (2002, 200) points out, "[o]ffensiveness is never an intrinsic quality of the word, but of the way it's used and its context" and thus each word can in theory be used non-offensively. On the other hand, Burridge (2004, 57) also controversially notes that the derogatory meaning of a word always dominates the way expressions are interpreted, and speakers are not often willing to take the risk of using a 'bad' word.

Even though Wachal (2002, 201) notes that ethnophaulisms have become more tabooed in dictionaries, this does not seem to apply to ethnophaulisms denoting all nationalities. For example, the words denoting Americans were marked as offensive in only 51% of all cases. Moreover, the entries that indicated offensiveness often had wordings such as *sometimes offensive*. In addition, words denoting the British were also often stated to be insulting only in some contexts, even though the general rate for the indications of offensiveness was higher than with words denoting Americans. At least ethnic slurs denoting these two nationalities did not seem to be taboos. Moreover, the label *taboo* or other labels indicating taboo status were not often used with any ethnic slurs of the study.

Words for other nationalities were treated as more offensive than words denoting the British and Americans. Even though the general rate for the indications of offensiveness for the British (76%) was higher than the rate for the French (70%), and the same as the rate for the Irish, entries for the words denoting the French and the Irish more often stated the words to be regularly offensive in each case and their offensiveness was more often emphasised. The rate for the indications of offensiveness for Germans was 77%. The words denoting Mexicans were marked as the most offensive (92%). The variation between the rates shows that Landau (2001, 232) has a point when arguing that in

dictionaries words denoting some nationalities are treated as more offensive than words denoting other nationalities.

Nevertheless, it should be taken into account that the dictionaries studied were not always unanimous when labelling words for the same nationality. Moreover, there was variation in the degree of offensiveness of words denoting the same nationality. The variation in the degree of offensiveness is natural: as Battistella (2005, 83) notes, words can be offensive in different ways. Moreover, as Saeed (2009, 65) argues, exact synonyms are very rare, if they exist at all. The economy of language would not tolerate the existence of two ethnophaulisms denoting the same nationality, being equally offensive and being identical in all other forms of meaning, too.

In the corpus analysis, the frequencies of ethnophaulisms corresponded quite well to the inclusion rates in the dictionaries. In dictionaries, the exclusion of some words was explained by dialectal factors, and this was also the case in the BNC. Other suggested reasons for the mismatch between the corpus and dictionary data were the estimations based on which the frequencies were formed, and in the case of Mexicans, the little need for the British to refer to Mexicans because of the limited contact between Britain and Mexico.

In the more detailed corpus analysis, it was stated that the labelling of *Pom* corresponded well to the corpus data. *Boche* seemed to be presented as a more offensive word in the dictionaries than in the BNC. This kind of policy is not surprising if ethnic slurs really have become more tabooed, as Wachal (2002, 201) notes. On the other hand, *Yank* seemed to be more disparagingly used in the BNC than it was indicated in the dictionaries examined. In addition, there were instances of all three words being used to convey a positive message. As Landau (2001, 233) points out, no term is insulting in all contexts.

9. Conclusion

This study examined the treatment of insulting nationality words in dictionaries and in the BNC. The aim of the study was to see how often and in what ways ethnophaulisms are categorised as offensive in dictionaries, what kind of differences and similarities there are between dictionary types and single volumes, and whether there are differences between the indications of offensiveness for different nationalities. The BNC was examined to see whether the inclusion and exclusion of certain terms in dictionaries corresponds to the corpus findings, and also to achieve a closer look of three ethnic slurs included in the study. The material of the study consisted of twenty dictionaries, thirty-seven words (denoting six nationalities) and a corpus.

In the dictionary analysis, it was found out that the ethnophaulisms examined were classified as offensive in 75% of all the entries studied. The words were most often marked as offensive in British learner's dictionaries, which indicated the insulting connotation in 98% of all entries and were thus almost uniform. British general purpose dictionaries marked 76% and American general purpose dictionaries 73% of all entries as offensive, but there was great variation between single volumes. American collegiate dictionaries marked 69% of all entries as offensive, and the rates were quite similar in all three collegiate dictionaries studied. Slang dictionaries were the group that used the fewest indications of offensiveness; they were present in 48% of cases.

Labels were the most common technique to indicate the offensiveness of the words studied. The second most common way was using a combination of different techniques. Sometimes the combination was used to inform the reader about both the offensive and the pejorative connotations that the word had, but often different techniques only repeated the same information. The third most common way of stating offensiveness was mentioning it in the definition, and usage notes were the least employed option. Only a minority of the dictionaries examined consistently used one technique to indicate offensiveness in all entries. Many dictionaries used at least two different ways to state the

negative connotation: sometimes they used one technique, sometimes the other. Furthermore, some unnecessary repetition was found in many entries.

The nationality which was most often classified as offensive was Mexicans (92%). The second highest rate was for Germans (77%), followed by the Irish and the British (76%), the French (72%), and finally Americans (51%). Words for Group 3 were classified as most offensive, followed by Group 2 and Group 1. It should be noted that mere percentages do not present the entire picture. Even though the words for the British and the Irish were marked as insulting equally often, the details revealed that the words for the Irish were in fact treated as more offensive; the offensiveness of these words was emphasised more often, and it was often stated in connection with the words denoting the British that those words are not offensive in all contexts. When examining the details, the nationality words seen as least offensive were the words for Americans and Brits.

Different dictionary groups did not treat the words denoting the same nationality unanimously very often; sometimes one dictionary group did not label any words denoting a particular nationality as offensive, whereas another dictionary group labelled all of the words denoting the same nationality as insulting. It should be noted that not all words denoting the same nationality were treated as equally offensive; in many cases, some words were stated to be remarkably more insulting than others. Several words were only included in a few dictionaries. Sometimes the inclusion or exclusion of certain words could be explained by the place of publication of the dictionary; American dictionaries listed more words for Mexicans, for example. There were many discrepancies between dictionaries when indicating words as offensive. Some dictionaries stated a certain word to be offensive, some mentioned that it can sometimes be insulting, some stated that the word is very rude and some did not state the offensiveness of that word at all. Furthermore, there were many discrepancies between the different volumes when they aimed at specifying the exact nature of the negative connotation: that is to say,

whether the word should be marked as offensive, pejorative or both. Of course, no total unanimity was expected but the variation was greater than had been anticipated.

The information given in the dictionaries corresponded quite well to the corpus data. In some cases, however, there were few or zero instances in the corpus of words listed in the majority of the dictionaries examined, and one case where a word rarely listed in the dictionaries studied was estimated to have several tokens in the BNC. Calculated estimations may have had an effect on the mismatch between the dictionary and corpus data, and the mismatch may also have been produced because of dialectal factors; the BNC represents British English and does not have as many words from other varieties of English as it has from British English. One reason for the non-corresponding corpus and dictionary data was also suggested to be the limited contact between Britain and the nationality referred to. With *bogtrotter*, however, no explanation for the mismatch was found.

A more detailed corpus analysis was executed with *Boche*, *Pom* and *Yank*. Based on the results, it was stated that the labelling of *Pom* corresponded best to the corpus data, as it was marked as offensive in some contexts, but not in all of them. *Boche* seemed to be presented as a more offensive word in the dictionaries than in the BNC, and *Yank* seemed to be more disparagingly used in the BNC than was suggested by the dictionaries examined.

There are many applications for the findings of the present study. The study illustrated that dictionaries are not always consistent when indicating that a particular word can be treated as offensive. A reader of the dictionary would benefit from a simple system where the negative connotations of the word were always stated in the same way, such as with a label or in the definition. Then it would be easier to see if a certain expression was politically correct or not, as the reader would always know where to look for the information. The markers could in places be simplified; if unnecessary repetition was removed, space would be saved. Moreover, it is clear that the distinction

into offensive and disparaging uses is difficult to make. Perhaps both the reader and the compiler of a dictionary would benefit from a system where a single marker would stand for both uses.

There are several possibilities for further study. Of course, the number of nationalities and dictionaries included in the study could easily be increased and thus the present study could be expanded. The words denoting the British could be examined more thoroughly by comparing the words denoting the English and the Scottish, for example. The dictionary analysis could be more detailed; the distinction into offensive, disparaging and taboo uses could be presented for every word examined. Another corpus, such as COCA, could also be added and the corpus analysis could be more detailed: a detailed corpus analysis could be executed on all words included in the dictionary analysis, as in the present study only three words were examined more thoroughly.

These are not, however, the only options. The study could also be expanded to the direction of historical lexicography, for example. Earlier editions of the dictionaries studied could be examined and the findings could be compared to the results of the present study. It would be interesting to examine when warnings for ethnic slurs have started to appear in dictionaries and whether different types of dictionaries have started to include these warnings simultaneously or during different periods of time. Historical corpora, such as COHA, could also be examined and it could be seen whether the development of ethnophaulisms into taboo terms would be visible in the corpus.

Moreover, the origins of ethnic slurs could be examined. It was stated in Section 2.4 that ethnic slurs usually develop when there is migration, immigration, war or other kinds of conflicts, for example. The origins of the ethnophaulisms examined in this thesis could be identified and the most common sources for ethnic slurs could be listed. It could also be examined if the sources for ethnic slurs are nowadays different from those of the older ethnophaulisms. The origin could be examined from another angle, too: it seems that many ethnic slurs derive from common sources, such as foods

that people belonging to a particular nationality are thought to eat (e.g. *frog, Kraut, sausage*), or word plays of rhyming slang, such as *septic*.

Much research remains to be done in the field of ethnophaulisms. I hope that I will be able to continue my studies in this area, for instance with some of the possible themes mentioned above.

References

A. Primary sources

Dictionaries and thesauri

Brookes, Ian et al., eds. 2014. *Collins English Dictionary.* 12th ed. 1st ed. 1874. Glasgow: HarperCollins Publishers.

Dalzell, Tom, and Terry Victor. 2008. *The Concise New Partridge Dictionary of Slang and Unconventional English.* 1st published as *The Dictionary of Slang and Unconventional English* in 1937. Abingdon and New York: Routledge.

Davidson, George, ed. 2002. *Roget's Thesaurus of English Words and Phrases.* 150th anniversary ed. 1st ed. 1852. London: the Penguin Group.

Flexner, Stuart Berg, ed. 2001. *Random House Webster's Unabridged Dictionary* (with New Words section). 2nd ed. 1st ed. 1966. New York: Random House, Inc.

Gove, Philip Babcock, ed. 1961 (with Addenda section, 1966–2002). *Webster's Third New International Dictionary, Unabridged.* 3rd ed. 1st ed. 1909. Springfield: Merriam-Webster Inc.

Kay, Christian et al., eds. 2009. *Historical Thesaurus of the Oxford English Dictionary*. Oxford: Oxford University Press.

Kleinedler, Steve, ed. 2014. *Webster's New World College Dictionary*. 5th ed. 1st ed. 1951. Boston and New York: Houghton Mifflin Harcourt Publishing Company.

Marr, Vivian, ed. 2011. *The Chambers Dictionary*. 12th ed. 1st ed. 1901. London: Chambers Harrap Publishers Ltd.

Mayor, Michael, ed. 2009. *Longman Dictionary of Contemporary English for Advanced Learners*. 5th ed. 1st ed. 1978. Essex: Pearson Education Limited.

McIntosh, Colin, ed. 2013. *Cambridge Advanced Learner's Dictionary*. 4th ed. 1st ed. 1995. Cambridge: Cambridge University Press.

Mish, Frederick C., ed. 2011 (principal copyright 2003). *Merriam-Webster's Collegiate Dictionary.* 11th ed. 1st ed. 1898. Springfield: Merriam-Webster, Incorporated.

Pickett, Joseph P., ed. 2011. *The American Heritage Dictionary of the English Language*. 5th ed. 1st ed. 1969. Boston and New York: Houghton Mifflin Harcourt Publishing Company.

Pickett, Joseph P., ed. 2007. *The American Heritage College Dictionary.* 4th ed. (date of 1st ed. not given). Boston and New York: Houghton Mifflin Company.

Proffitt, Michael, ed. 2015. *The Oxford English Dictionary* online. 3rd ed. (in preparation). 1st ed. 1884. Oxford: Oxford University Press. Available from: http://oed.com/ [Accessed 4 February 2015]

Rooney, Kathy, ed. 2004. *Bloomsbury English Dictionary.* 2nd ed. 1st ed. 1999. London: Bloomsbury Publishing Plc.

Rooney, Kathy, ed. 1999. *Encarta World English Dictionary*. 1st ed. London: Bloomsbury Publishing Plc.

Rooney, Kathy, ed. 1999. *Encarta World English Dictionary*. 1st ed. New York: St. Martin's Press.

Rundell, Michael, ed. 2014. *Macmillan English Dictionary* online. 1st ed. 2002. London: Macmillan Publishers Limited. Available from: http://www.macmillandictionary.com/ [Accessed 3 June 2014]

Sinclair, John et al., eds. 2014. *Collins COBUILD Advanced Learner's Dictionary*. 8th ed. 1st ed. 1987. Glasgow: HarperCollins Publishers.

Stevenson, Angus, ed. 2010. *Oxford Dictionary of English.* 3rd ed. 1st ed. 1998. Oxford: Oxford University Press.

Thorne, Tony. 2007. *Dictionary of Contemporary Slang.* 3rd ed. 1st ed. 1990. London: A & C Black Publishers Ltd.

Turnbull, Joanna, ed. *Oxford Advanced Learner's Dictionary.* 2010. 8th ed. 1st ed. 1948. Oxford: Oxford University Press.

Corpora and other online sources

The British National Corpus (CQP-edition, version 4.3). Accessed using the BNCweb interface at https://bncweb.uta.fi [Accessed 8–9 August 2014]

The Racial Slur Database. 1999–. Available from: http://www.rsdb.org/ [Accessed 10 January 2014]

B. Works cited

Adams, Michael. 2009. *Slang: The People's Poetry.* New York: Oxford University Press, Inc.

Aguila, Emma et al. 2012. *United States and Mexico: Ties That Bind, Issues That Divide.* Santa Monica: RAND Corporation.

Allan, Keith, and Kate Burridge. 2006. *Forbidden Words: Taboo and the Censoring of Language.* Cambridge: Cambridge University Press.

Atkins, B. T. Sue, and Michael Rundell. 2008. *Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.

Battistella, Edwin L. 2005. *Bad Language: Are Some Words Better Than Others?* Oxford: Oxford University Press.

Bauer, Laurie. 1994. *Watching English Change: An Introduction to the Study of Linguistic Change in Standard Englishes in the Twentieth Century.* London: Longman Group Limited.

Béjoint, Henri. 2010. *Lexicography of English: from Origins to Present*. Oxford: Oxford University Press.

Brewer, Charlotte. 2005. "Authority and Personality in the Oxford English Dictionary". *Transactions of the Philological Society* 103, 3: 261–301.

Burridge, Kate. 2002. *Blooming English: Observations on the Roots, Cultivation and Hybrids of the English language.* Cambridge: Cambridge University Press.

Burridge, Kate. 2004. *Weeds in the Garden of Words: Further Observations on the Tangled History of the English Language.* Cambridge: Cambridge University Press.

Campbell, Elisabeth. 2006. "Le traitement des insultes sexistes et racistes dans les dictionnaires bilingues". *The French Review* 80, 1: 112–137.

Cann, Ronnie. 2011. "Sense Relations". In *Semantics: An International Handbook of Natural Language Meaning,* Vol. 1, edited by Claudia Maienborn, Klaus von Heusinger, and Paul Portner, pp. 456–479. Berlin: Walter de Gruyter GmbH & Co.

Coakley, Maurice. 2012. *Ireland in the World Order: A History of Uneven Development*. London: Pluto Press.

Croom, Adam M. 2013. "How to Do Things with Slurs: Studies in the Way of Derogatory Words". *Language & Communication* 33, 3: 177–204.

Cruse, D. A. 1986. *Lexical Semantics.* Cambridge: Cambridge University Press.

Cruse, Alan. 2011. *Meaning in Language: An Introduction to Semantics and Pragmatics.* 3rd ed. 1st ed. 1999. Oxford: Oxford University Press.

Hanks, Patrick. 2000. "Do Word Meanings Exist?" In *Practical Lexicography: A Reader,* edited by Thierry Fontenelle, pp. 125–134. Oxford: Oxford University Press.

Harley, Heidi. 2006. *English Words: A Linguistic Introduction.* Oxford: Blackwell.

Henderson, Anita. 2003. "What's in a Slur?" *American Speech* 78, 1: 52–74.

Herbst, Thomas. 2010. *English Linguistics: A Coursebook for Students of English.* Berlin and New York: De Gruyter Mouton.

Hudson, Anthony. 2003. "Fighting Words". *Index on Censorship* 32, 4: 45–51.

Hughes, Geoffrey. 2006. *Encyclopedia of Swearing: The Social History of Oaths, Profanity, Foul Language, and Ethnic Slurs in the English-Speaking World*. Armonk: M. E. Sharpe, Inc.

Jackson, Howard. 1988. *Words and Their Meaning.* London: Longman.

Jackson, Howard. 2002. *Lexicography: An Introduction.* London: Routledge.

Jaszczolt, K. M. 2002. *Semantics and Pragmatics: Meaning in Language and Discourse.* Harlow: Pearson Education Limited.

Julios, Christina. 2008. *Contemporary British Identity: English Language, Migrants and Public Discourse.* Aldershot: Ashgate.

Kearns, Kate. 2000. *Semantics.* London: Macmillan Press Ltd.

Kennedy, Randall. 2002. *Nigger: The Strange Career of a Troublesome Word.* New York: Vintage Books.

Landau, Sidney I. 2001. *Dictionaries: the Art and Craft of Lexicography*. 2nd ed. Cambridge: Cambridge University Press.

Lillo, Antonio. 2001. "From Alsatian Dog to Wooden Shoe: Linguistic Xenophobia in Rhyming Slang". *English Studies: A Journal of English Language and Literature* 82, 4: 336–348.

Leith, Murray Stewart, and Daniel Soule. 2011. *Political Discourse and National Identity in Scotland.* Edinburgh: Edinburgh University Press.

Loomba, Ania. 1998. *Colonialism/Postcolonialism.* London: Routledge.

Lyons, John. 1977. *Semantics. Volume I.* Cambridge: Cambridge University Press.

Lyons, John. 1995. *Linguistic Semantics.* Cambridge: Cambridge University Press.

Murphy, M. Lynne. 2010. *Lexical Meaning.* Cambridge: Cambridge University Press.

Norri, Juhani. 2000. "Labelling of Derogatory Words in Some British and American Dictionaries". *International Journal of Lexicography* 13, 2: 71–106.

Peters, Pam, and Kate Burridge. 2012. "English in Australia and New Zealand". In *Topics in English Linguistics 80: Areal Features of the Anglophone World*, edited by Raymond Hickey, Bernd Kortmann, and Elizabeth Closs Traugott, pp. 233–258. Berlin/Boston: Walter de Gruyter GmbH.

Ptaszynski, Marcin Overgaard. 2010. "Theoretical Considerations for the Improvement of Usage Labelling in Dictionaries: a Combined Formal-Functional Approach". *International Journal of Lexicography* 23, 4: 411–442.

Rice, Diana R. et al. 2010. "What Did You Just Call Me? European and American Ratings of the Valence of Ethnophaulisms". *Journal of Language and Social Psychology* 29, 1: 117–131. (first published on December 16, 2009).

Roback, A. A. 1944. *Dictionary of International Slurs.* Cambridge, MA: Sci-Art Publishers.

Saeed, John I. 2009. *Semantics*. Chichester: John Wiley & Sons Ltd.

Verkuyten, Maykel, and Jochem Thijs. 2010. "Ethnic Minority Labeling, Multiculturalism, and the Attitude of Majority Group Members". *Journal of Language and Social Psychology* 29, 4: 467–477.

Wachal, Robert S. 2002. "Taboo or Not Taboo: That Is the Question". *American Speech* 77, 2: 195–206.

Online sources

Burnard, Lou. 2010. "What is the BNC?" Oxford, University of Oxford. Available from http://www.natcorp.ox.ac.uk/corpus/index.xml [Accessed 6 August 2014]

"Scottish Referendum: Scotland Votes 'No' to Independence". 2014. BBC online. Available from: http://www.bbc.com/news/uk-scotland-29270441 [Accessed 6 February 2015]

Appendices

Appendix 1. Tables for the indications of offensiveness

Table 13. Frequencies of indications of offensiveness (the British)

		Words included	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these:
Br. general purpose dict.	*BED*	5	0	3	2	0	0	5
	CED	7	4	0	3	0	0	3
	ChD	7	3	3	1	0	0	4
	EWEDL	5	0	2	0	0	3	5
	ODE	5	1	4	0	0	0	4
	OED	7	2	3	2	0	0	5
	TOTAL	**36**	**10 (28 %)**	**15**	**8**	**0**	**3**	**26 (72%)**
Am. general purpose dict.	*AHD*	2	1	0	0	0	1	1
	EWEDNY	3	0	2	0	0	1	3
	RHD	4	1	3	0	0	0	3
	W3	6	3	0	0	3	0	3
	TOTAL	**15**	**5 (33%)**	**5**	**0**	**3**	**2**	**10 (67%)**
Br. learner's dict.	*CALD*	5	1	3	0	1	0	4
	COBUILD	3	0	0	0	1	2	3
	LDCE	5	0	0	2	1	2	5
	MED	5	0	0	2	1	2	5
	OALD	5	0	3	2	0	0	5
	TOTAL	**23**	**1 (4%)**	**6**	**6**	**4**	**6**	**22 (96%)**
Am. collegiate dict.	*AHCD*	2	1	0	0	0	1	1
	MWCD	3	1	2	0	0	0	2
	WNW	3	1	0	0	2	0	2
	TOTAL	**8**	**3 (38%)**	**2**	**0**	**2**	**1**	**5 (63%)**
Slang dict.	*DCS*	4	1	0	3	0	0	3
	DSUE	4	2	0	2	0	0	2
	TOTAL	**8**	**3 (38%)**	**0**	**5**	**0**	**0**	**5 (63%)**
ALL DICT.	**TOTAL**	**90**	**22 (24%)**	**28**	**19**	**9**	**12**	**68 (76%)**

Table 14. Frequencies of indications of offensiveness (Americans)

		Words included (out of 5 words)	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these:
Br. general purpose dict.	*BED*	5	0	2	2	0	1	5
	CED	4	2	2	0	0	0	2
	ChD	3	3	0	0	0	0	0
	EWEDL	3	0	1	2	0	0	3
	ODE	3	0	3	0	0	0	3
	OED	4	3	1	0	0	0	1
	TOTAL	**22**	**8 (36%)**	**9**	**4**	**0**	**1**	**14 (64%)**
Am. general purpose dict.	*AHD*	2	2	0	0	0	0	0
	EWEDNY	3	0	2	1	0	0	3
	RHD	3	3	0	0	0	0	0
	W3	3	3	0	0	0	0	0
	TOTAL	**11**	**8 (73%)**	**2**	**1**	**0**	**0**	**3 (27%)**
Br. learner's dict.	*CALD*	2	0	2	0	0	0	2
	COBUILD	2	0	0	0	2	0	2
	LDCE	2	0	0	0	2	0	2
	MED	2	0	0	0	1	1	2
	OALD	2	0	0	0	0	2	2
	TOTAL	**10**	**0 (0%)**	**2**	**0**	**5**	**3**	**10 (100%)**
Am. collegiate dict.	*AHCD*	2	2	0	0	0	0	0
	MWCD	3	3	0	0	0	0	0
	WNW	3	3	0	0	0	0	0
	TOTAL	**8**	**8 (100%)**	**0**	**0**	**0**	**0**	**0 (0%)**
Slang dict.	*DCS*	3	3	0	0	0	0	0
	DSUE	3	1	0	1	1	0	2
	TOTAL	**6**	**4 (67%)**	**0**	**1**	**1**	**0**	**2 (33%)**
ALL DICT	**TOTAL**	**57**	**28**	**13**	**6**	**6**	**4**	**29 (51%)**

Table 15. Frequencies of indications of offensiveness (the Irish)

		Words included (out of 6 words)	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these:
Br. general purpose dict.	*BED*	4	0	0	2	0	2	4
	CED	4	1	1	2	0	0	3
	ChD	5	1	3	1	0	0	4
	EWEDL	4	0	0	0	0	4	4
	ODE	4	0	4	0	0	0	4
	OED	6	3	2	0	1	0	3
	TOTAL	**27**	**5 (19%)**	**10**	**5**	**1**	**6**	**22 (81%)**
Am. general purpose dict.	*AHD*	3	0	0	0	0	3	3
	EWEDNY	4	0	0	0	0	4	4
	RHD	4	0	4	0	0	0	4
	W3	6	3	0	0	3	0	3
	TOTAL	**17**	**3 (18%)**	**4**	**0**	**3**	**7**	**14 (82%)**
Br. learner's dict.	*CALD*	1	0	0	0	0	1	1
	COBUILD	0	0	0	0	0	0	0
	LDCE	2	0	0	0	0	2	2
	MED	2	0	0	0	0	2	2
	OALD	2	0	0	1	0	1	2
	TOTAL	**7**	**0 (0%)**	**0**	**1**	**0**	**6**	**7 (100%)**
Am. collegiate dict.	*AHCD*	3	0	0	0	0	3	3
	MWCD	2	0	2	0	0	0	2
	WNW	2	0	0	0	2	0	2
	TOTAL	**7**	**0 (0%)**	**2**	**0**	**2**	**3**	**7 (100%)**
Slang dict.	*DCS*	3	2	0	1	0	0	1
	DSUE	6	6	0	0	0	0	0
	TOTAL	**9**	**8 (89%)**	**0**	**1**	**0**	**0**	**1 (11%)**
ALL DICT.	**TOTAL**	**67**	**16**	**16**	**7**	**6**	**22**	**51 (76%)**

Table 16. Frequencies of indications of offensiveness (the French)

		Words included (out of 7 words)	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these:
Br. general purpose dict.	*BED*	1	0	0	1	0	0	1
	CED	4	2	0	2	0	0	2
	ChD	6	2	0	4	0	0	4
	EWEDL	1	0	0	0	0	1	1
	ODE	3	0	3	0	0	0	3
	OED	7	0	7	0	0	0	7
	TOTAL	**22**	**4 (18%)**	**10**	**7**	**0**	**1**	**18 (82%)**
Am. general purpose dict.	*AHD*	1	0	0	0	0	1	1
	EWEDNY	1	0	0	0	0	1	1
	RHD	2	1	1	0	0	0	1
	W3	5	2	0	0	3	0	3
	TOTAL	**9**	**3 (33%)**	**1**	**0**	**3**	**2**	**6 (67%)**
Br. learner's dict.	*CALD*	1	0	1	0	0	0	1
	COBUILD	1	0	0	0	1	0	1
	LDCE	1	0	0	0	0	1	1
	MED	1	0	0	0	0	1	1
	OALD	1	0	0	1	0	0	1
	TOTAL	**5**	**0 (0%)**	**1**	**1**	**1**	**2**	**5 (100%)**
Am. collegiate dict.	*AHCD*	1	0	0	0	0	1	1
	MWCD	1	0	1	0	0	0	1
	WNW	2	1	0	0	1	0	1
	TOTAL	**4**	**1 (25%)**	**1**	**0**	**1**	**1**	**3 (75%)**
Slang dict.	*DCS*	2	2	0	0	0	0	0
	DSUE	4	3	0	1	0	0	1
	TOTAL	**6**	**5 (83%)**	**0**	**1**	**0**	**0**	**1 (17%)**
ALL DICT.	**TOTAL**	**46**	**13**	**13**	**9**	**5**	**6**	**33 (72%)**

Table 17. Frequencies of indications of offensiveness (Germans)

		Words included (out of 6 words)	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these:
Br. general purpose dict.	*BED*	3	0	0	2	0	1	3
	CED	3	1	1	1	0	0	2
	ChD	3	1	2	0	0	0	2
	EWEDL	3	0	0	0	0	3	3
	ODE	4	2	2	0	0	0	2
	OED	6	1	2	0	0	3	5
	TOTAL	**22**	**5 (23%)**	**7**	**3**	**0**	**7**	**17 (77%)**
Am. general purpose dict.	*AHD*	3	1	1	0	0	1	2
	EWEDNY	3	0	0	0	0	3	3
	RHD	5	1	4	0	0	0	4
	W3	6	1	0	0	5	0	5
	TOTAL	**17**	**3 (18%)**	**5**	**0**	**5**	**4**	**14 (82%)**
Br. learner's dict.	*CALD*	0	0	0	0	0	0	0
	COBUILD	0	0	0	0	0	0	0
	LDCE	1	0	0	0	0	1	1
	MED	2	0	0	0	0	2	2
	OALD	2	0	0	0	0	2	2
	TOTAL	**5**	**0 (0%)**	**0**	**0**	**0**	**5**	**5 (100%)**
Am. collegiate dict.	*AHCD*	3	1	0	0	0	2	2
	MWCD	2	1	1	0	0	0	1
	WNW	2	0	0	0	2	0	2
	TOTAL	**7**	**2 (29%)**	**1**	**0**	**2**	**2**	**5 (71%)**
Slang dict.	*DCS*	1	1	0	0	0	0	0
	DSUE	4	2	0	2	0	0	2
	TOTAL	**5**	**3 (60%)**	**0**	**2**	**0**	**0**	**2 (40%)**
ALL DICT.	**TOTAL**	**56**	**13**	**13**	**5**	**7**	**18**	**43 (77%)**

Table 18. Frequencies of indications of offensiveness (Mexicans)

		Words included (out of 6 words)	Off. not stated	Off. stated with a label	Off. stated in the definition	Off. stated with a usage note	Multiple indications for off.	Off. stated with a label, usage note, in the definition or with a combination of these:
Br. general purpose dict.	*BED*	2	0	0	0	0	2	2
	CED	2	0	1	1	0	0	2
	ChD	4	1	3	0	0	0	3
	EWEDL	1	0	0	0	0	1	1
	ODE	5	0	5	0	0	0	5
	OED	5	2	1	2	0	0	3
	TOTAL	**19**	**3 (16%)**	**10**	**3**	**0**	**3**	**16 (84%)**
Am. general purpose dict.	*AHD*	4	0	0	0	0	4	4
	EWEDNY	4	0	0	0	0	4	4
	RHD	4	0	4	0	0	0	4
	W3	4	1	0	0	3	0	3
	TOTAL	**16**	**1 (6%)**	**4**	**0**	**3**	**8**	**15 (94%)**
Br. learner's dict.	*CALD*	1	0	1	0	0	0	1
	COBUILD	1	0	1	0	0	0	1
	LDCE	2	0	0	0	0	2	2
	MED	1	0	0	0	0	1	1
	OALD	3	0	0	0	0	3	3
	TOTAL	**8**	**0 (0%)**	**2**	**0**	**0**	**6**	**8 (100%)**
Am. collegiate dict.	*AHCD*	3	0	0	0	0	3	3
	MWCD	4	0	4	0	0	0	4
	WNW	4	0	0	0	4	0	4
	TOTAL	**11**	**0 (0%)**	**4**	**0**	**4**	**3**	**11 (100%)**
Slang dict.	*DCS*	6	1	0	5	0	0	5
	DSUE	6	0	0	3	3	0	6
	TOTAL	**12**	**1 (8%)**	**0**	**8**	**3**	**0**	**11 (92%)**
ALL DICT.	**TOTAL**	**66**	**5**	**20**	**11**	**10**	**20**	**61 (92%)**

Appendix 2. Calculations for the corpus estimations

Table 19. Calculations for the corpus estimations (the answers have been rounded to full figures)

Word	Calculation
Fritz	0,01x208=2
frog	0,05x938=47
harp	0x300=0
Jerry	0,03x529=16
Jock	0,05x246=12
Mick	0,01x1496=15
paddy	0x899=0
Pat	0x1917=0
sausage	0,01x965=10
Taff(y)	0x171=0
Yank	0,85x112=95
Yankee	0,19=142=27