Treatment of Controversial Topics
in English Learner Dictionaries,
with Focus on the *Oxford Advanced Learner's Dictionary*

Paula Antila
University of Tampere
School of Modern Languages and Translation Studies
English Philology
Pro Gradu Thesis
May 2008

i

Tiivistelmä (Abstract)

Tämän työn tarkoituksena on tutkia ns. kiisteltyjen puheenaiheiden ja niihin liittyvien sanojen käsittelyä englanninkielisissä oppijasanakirjoissa. Kiistellyillä puheenaiheilla tarkoitetaan tässä työssä aiheita, jotka synnyttävät kielenpuhujien keskuudessa hyvin voimakkaita erisuuntaisia tunteita. Tutkimuskohdetta lähestytään tabu-käsitteen avulla.

Sanojen käsittelyllä viitataan sekä niiden olemassaoloon sanakirjassa että informaatioon sanojen tyylillisistä eroista ja merkitysvivahde-eroista. Tämä informaatio on useimmiten merkittynä sanakirja-artikkeliin erillisenä kielenkäyttömerkintänä (engl. *usage label*). Kielteisestä sävystä tai merkitysvivahteesta varoittavat merkinnät ovat tämän tutkimuksen ydinosa.

Työssä tutkittiin erään oppijasanakirjasarjan kehitystä kiisteltyjen sanojen käsittelyn suhteen. Tutkimuskohteena on *Oxford Advanced Learner's Dictionary* –sanakirjasarja, josta tutkittiin ensimmäiset kuusi painosta lähtien vuodesta 1948 vuoteen 2000.

Materiaalina oli 272 sanaa, jotka kuuluivat kahdeksaan erilliseen aihealueeseen. Tutkitut aihealueet olivat prostituoitu, homoseksuaali, seksuaalinen kanssakäyminen, pakarat, vihaa ilmaisevat huudahdukset, etniseen vähemmistöön kuuluva henkilö, typerys sekä lihavuutta kuvaavat adjektiivit. Sanat etsittiin kustakin sanakirjasta ja löytyneet kielenkäyttömerkinnät huomioitiin. Lisäksi tutkittiin kunkin sanakirjan käyttäjäoppaan antama informaatio erilaisten aihealueiden sisällyttämisestä kyseiseen teokseen sekä informaatio kielenkäyttömerkinnöistä.

Tutkimuksessa todettiin, että kokonaisuutena sanakirjat ovat kehittyneet systemaattisemman kielenkuvauksen suuntaan. Uudemmat sanakirjat luettelevat useampia kiisteltyjä sanoja kuin vanhemmat teokset. Kielenkäyttömerkinnät ovat samaten tulleet järjestelmällisemmiksi. Uudemmat sanakirjat esittelevät merkintäjärjestelmänsä käyttäjäoppaissaan ja selittävät, mitä kullakin merkillä tarkoitetaan. Suurin edistysaskel otettiin neljännen painoksen myötä (ilmestymisvuosi 1989), jolloin käyttäjäoppaan linjoja alettiin noudattaa myös käytännössä.

Kehitys ei kuitenkaan ole ollut samanlaista kaikilla aihealueilla. Jotkut sanaryhmät ovat sanakirjoissa paremmin edustettuina kuin toiset, esimerkkinä kattavasti käsitellyistä aiheista lihavuutta kuvaavat sanat. Toisaalta, jotkut sanaryhmät ovat suhteellisesti paremmin edustettuina vanhemmissa kuin uudemmissa teoksissa (prostituoitu, etniseen vähemmistöön kuuluva henkilö, typerys) ja päinvastoin (homoseksuaali, pakarat). Samaten kielenkäyttömerkintöjen yleisyys vaihtelee sanaryhmittäin.

Mikään yksittäinen ilmiö ei selitä näitä eroja täydellisesti. Norrin (1999) huomion erilaisten tabutyyppien käsittelystä sanakirjoissa havaittiin pitävän pääosin paikkansa myös tämän

tutkimuksen materiaalin suhteen. Joidenkin puheenaiheiden esiin nostaminen rikkoo hyviä tapoja vastaan (engl. *taboos of propriety*), kun taas toisia puheenaiheita vältetään niiden arkaluontoisuuden takia (engl. *taboos of delicacy*). Norrin (1999) mukaan ensimmäiseen ryhmään kuuluvat sanat saavat toisia todennäköisemmin varoittavan merkinnän.

Toisinaan sanan merkitys itsessään paljastaa kielteisen sävyn, jolloin käyttömerkintä on tarpeeton (Norri 2000). Näiden selittävien tekijöiden lisäksi havaittiin, että tässä työssä sanaryhmät edustivat erilaisia kiistanalaisuuden tasoja. Toiset aiheet olivat luonteeltaan vähemmän kiisteltyjä kuin toiset, mikä osaltaan vaikutti sekä sanojen esiintymiseen että käyttömerkintöihin.

Avainsanat: *Oxford Advanced Learner's Dictionary*, controversial topic, taboo, usage label

Contents

Abbreviations and Typographical Conventions

EFL   English as a foreign language

OALD1   *A Learner's Dictionary of Current English*, 1st edition.

OALD2   *The Advanced Learner's Dictionary of Current English*, 2nd edition.

OALD3   *Oxford Advanced Learner's Dictionary of Current English*, 3rd edition.

OALD4   *Oxford Advanced Learner's Dictionary of Current English*, 4th edition.

OALD5   *Oxford Advanced Learner's Dictionary of Current English*, 5th edition.

OALD6   *Oxford Advanced Learner's Dictionary of Current English*, 6th edition.

⚠   The symbol generally known as a sign of danger is used in this study similarly as in *OALD3-6* to stand for the sense 'taboo', as well as to replace the sign ⬠ used in *OALD2* to indicate 'vulgar or incorrect' uses.

1. Introduction

Sets of seemingly synonymous words are abundant in the English vocabulary. To name but one set, the *Oxford Thesaurus* (Urdang 1991 (1997-99)) lists the following thirty expressions as possible equivalents to the noun *girl*: *female*, *woman*, *lass*, *(young) lady*, *miss*, *mademoiselle*, *wench*, *Fräulein*, *maid*, *maiden*, *damsel*, *demoiselle*, *colleen*, *sheila*, *chick*, *filly*, *gal*, *bird*, *frail*, *skirt*, *piece*, *mouse*, *crumpet*, *bit of skirt* or *stuff*, *popsy*, *dame*, *broad*, *(bit of) San Quentin quail* or *jail-bait*. While native speakers usually have no difficulty in choosing the most fitting word for each situation, foreign learners may need guidance in recognizing semantic and stylistic differences between words. The form of a word nor its denotative meaning does not necessarily explain its status in the vocabulary. Using a word that is perhaps not the aptest one is, of course, usually not a gross mistake for a language learner. In a more unfortunate situation, however, an ill-chosen word may lead to negative effects such as misunderstanding or feelings of offence and awkwardness. All in all, appropriate wording is an essential factor in successful communication, as Trask wittily observes, "A foreign learner of English who talks solemnly about a fiddle concerto is not going to achieve the desired effect" (1995: 52).

The theme of this pro gradu thesis is the treatment of controversial topics and the relating vocabulary in English learner dictionaries. The term *controversial vocabulary* is adopted from Robert Burchfield who uses it in the meaning "the vocabulary that lies on or near the admission/exclusion boundary [of a dictionary]" (Burchfield 1989: 83). Thus Burchfield's use of the term covers quite a wide range of words from ones that are controversial for social and stylistic reasons (e.g. racial and religious words) to ones that are controversial for practical reasons connected with dictionary-making (e.g. various types of names, rare words and invented words). The focus of this thesis lies on topics and words that can be characterized as socially controversial; they are apt to provoke mixed reactions in people. The phenomenon of controversy will be approached in terms of taboo.

The aim of this thesis is to examine the development of one dictionary family as to the treatment of controversial topics and vocabulary from the first edition (*A Learner's Dictionary of Current English* (1948), hereafter referred to as *OALD1*) to the sixth edition (*Oxford Advanced Learner's Dictionary of Current English* (2000), hereafter referred to as *OALD6*). Changes manifested in each dictionary edition may, obviously, derive from changes in the language and the society as well as from changes in the dictionary making habits.

The term *treatment* is used here to cover two things: first, the inclusion of a word in a dictionary and second, the information which points out stylistic and social differences between words with similar meanings. This kind of information often appears in the form of usage labels in dictionaries, but relevant information can also be found in the definition proper or in the form of separate usage notes. For the sake of convenience, in this thesis the term *usage label* is employed in a general sense to refer to all forms of sociolinguistic information found in dictionaries. Usage labels are, of course, used to specify also other restrictions than stylistic and social ones: for instance, regional or temporal restrictions that a given word might have. These usage labels as well as all other types of information found in a dictionary entry – such as explanations of the meaning, spelling and pronunciation, syntactical information, and so on – lie outside the area of this study.

The idea for this type of diachronic study was raised in Norri's article on the labelling of derogatory words in contemporary dictionaries (2000), and this thesis follows Norri's study in many respects. Similarities and differences between this thesis and Norri's study will be detailed in the methods section (Chapter 6).

The following questions are formulated to clarify the research task further:

1. How has the inclusion of words of controversial topics changed from *OALD1* to *OALD6*?

2. How has the labelling of words of controversial topics changed from *OALD1* to *OALD6*?

Answers to these questions are sought by investigating a number of dictionary entries in each edition of *OALD*. The sample material consists of eight sets of words in which the members share a meaning feature. Often the words stand in a synonymous relationship to each other, but it would be too much to say that all the words within a given group are interchangeable with each other. The aim is to form a picture of how the semantic fields involving such controversial topics are treated as a whole, not to study the treatment of the stigmatized parts of those fields only. Thus, the sample material also contains items that have perhaps never been regarded as taboo. In addition to the sample material, the prefaces and user guides of the dictionaries are consulted.

As to the methods of the analysis, the investigation involves both quantitative and qualitative aspects. The objective of the study is to examine how the marking of negative connotation in controversial vocabulary has changed, not to evaluate the accuracy of the marking. In fact, it is not possible to make any reliable conclusions about the accuracy without a reference to material representing language as it was used at the respective points in time. Such a task is far beyond the scope of this study.

Chapter 2 deals with language variation and dictionary labelling as a device for indicating aspects of language variation on the lexical level. In Chapter 3, I will present two studies that are of direct relevance to this thesis. The first of these is Norri's study which gave the original idea and largely also the methods to this thesis, as already pointed out. The second study shares the interest in treatment of controversial topics in dictionaries. Chapter 4 concentrates on the concept of taboo and explains why certain topics are controversial. Chapters 5 and 6 contain details of the material and methods of this thesis. Chapter 7 is the first part of the analysis section: it explores what the *OALD* user guides tell about the treatment of controversial topics. Chapter 8, then, is the main chapter of the analysis, reporting the treatment of controversial topics in practice. The results of the eight semantic fields are amalgamated in Chapter 9. Concluding remarks are made in Chapter 10.

This thesis is essentially a diachronic study about lexicographical conventions. As the focus of the study lies on that part of the dictionary entry which specifies the stylistic and social

status of the entry word, this thesis involves also aspects of sociolinguistics – language variation and the concept of taboo. Ilson (1986: 127) calls this type of comparison of different editions of the same dictionary as *lexicographic archaeology* and points out the value of such studies: by measuring consecutive editions of one dictionary against each other, we may gain insights into the language as well as into the difficulties of making dictionaries. Ilson (1986: 127, 134) concludes that such a task can be useful in the training of people working in the field, as "understanding of past dictionaries can contribute to the improvement of future ones" (1986: 134).

Besides the professional, or academic, motives for the study, there are more personal viewpoints to this study, as well. My interest in the topic originates from experiences as both a learner and a teacher of English as a foreign language. Countless times have I turned to a dictionary when feeling unconfident about the best choice of words, and indeed quite often a dictionary has proved itself helpful. Many times have I also been wondering what kind of blunders do young language learners make when they use the rough language constantly heard on television. I would like to draw EFL teachers' attention to the key role they have in raising their students' language awareness. If dictionaries are what they, according to my view, should be, they may serve as a valuable aid to the teacher; an aid whose ability to instruct may extend beyond the classroom walls.

2. Usage Labels as a Reflection of Linguistic Variation

2.1 Dimensions of Linguistic Variation

Human language is not a mechanical sign system similar to, for instance, road signs in which each sign has only one interpretation and conversely, one particular meaning can be conveyed by means of one sign only. Rather, in natural human language a message can be put into words in various ways; it is said that language varies.

Quirk et al. (1985: 16) distinguish five main factors, or types of variation, which are present in every instance of language use and affect the format of a message: region, social group, field of discourse, medium and attitude. They state that the first two factors, regional and social background of a speaker or writer, affect the way in which a person speaks or writes in comparison to other people. The latter three terms – field of discourse, medium and attitude – refer to factors that make peoples' speech and writing differ according to the situation in which language is used. Thus Quirk et al. adopt a *user* versus *use* distinction in the factors of linguistic variation (1985: 16).

Types of variation according to region and social group quite obviously refer to a person's background: place of living on the one hand, and education, socioeconomic and ethnic group on the other. In addition, age and sex are said to influence a person's language use. The term *field of discourse* describes the type of activity which is being communicated, such as cookery, legal matters or machinery repairing. Variation according to the medium refers to differences that arise from whether the message is expressed in speech or in writing. Lastly, Quirk et al. identify varieties according to attitude. They observe that the speaker's or writer's attitude to the addressee, to the topic and to the purpose of communication affect the shape of a message. (Quirk et al 1985: 16-27)

In addition to the five major types of varieties, Quirk et al. name two sources of variation that are of lesser importance to their ends. The first of these is variation in time, but as their work concerns contemporary English, it is left without further attention (1985:1 16). In a diachronic study such as this thesis, variation in time is naturally relevant. The second minor type of variety is

variation according to interference (1985: 27). This type of variation signifies the influence of other languages in the speech of those who speak English as a second or foreign language.

2.2 Linguistic Variation at the Lexical Level

Hughes (1988:17-19) explains linguistic variation at the lexical level with the concept of *semantic field*. All expressions that refer to a given thing or topic are said to form one semantic field:
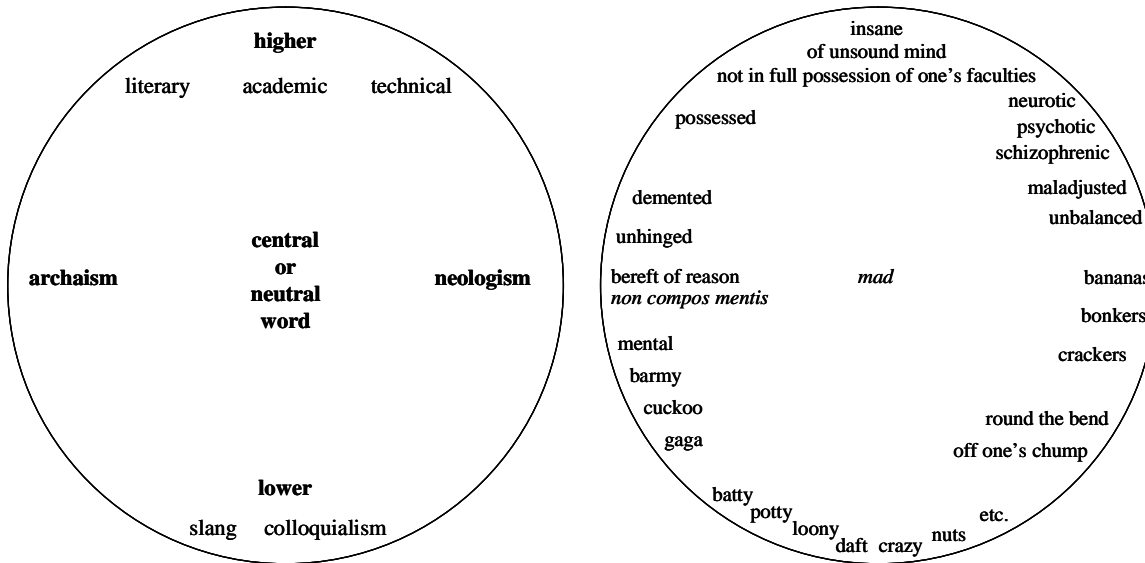
Figure 1 Registers in a semantic field and the semantic field of *mad* (adopted from Hughes 1988: 18-19).

A semantic field thus comprises different ways to express the same idea: the speaker or writer can choose, for instance, between a new word and an older one, between a literary expression and a more colloquial one, or between a blunt word and a politer one (Hughes 1988: 17-18). Hughes (1988: 19-20) also points out that not all semantic fields are as evenly balanced as the semantic field of *mad*. Some semantic fields are by their nature "bottom-heavy" or "top-heavy" (Hughes 1988: 19-20). In other words, some semantic fields contain more expressions that belong rather to the lower side of register[1] levels than to the higher side, or vice versa.

---

[1] Hughes is a proponent of the term *register* which others, like Quirk et al., have rejected. Hughes uses the term in the sense "language variation according to social role or social situation" (1988:17).

Allan and Burridge (1991: 29-30) call such groups of alternative expressions as *cross-varietal synonyms*. Such alternatives have the same denotation but differ in connotation; therefore, the expressions usually belong to different contexts, the writers point out (1991: 30). However, there is actually a two-way relationship between individual expressions and the style of a text as a whole: on the one hand, individual expressions make up the style of a text, and the intended style directs the choice of words, on the other (Allan and Burridge 1991: 31).

Palmer (1981: 92) rejects the claim that synonyms differ only in their connotations as inaccurate. First, even though Palmer sees word meaning as a composition of two parts – *cognitive meaning* and *emotive or evaluative meaning* – terms that he uses in quite a similar way that some writers use *denotation* and *connotation*, he says that these meaning parts are not distinctly separable from each other. Secondly, the term *connotation* has been used in various senses in the past. It is used to refer to such differences between expressions that actually are stylistic or dialectal differences, or cognitive differences between near-synonyms, Palmer (1981: 92) argues. Yet another sense in which the term appears are associations attached to a given word, which originate from the characteristics of the referent. For example, Palmer (1981: 92) explains that pigs are often thought to be dirty; thus the word *pig* has the connotation 'dirty'.

Palmer proposes that synonyms differ in at least five ways, some of which already were mentioned above (1981: 89-91): First, they belong to different dialects; second, they are typically used in different styles; third, they differ in their emotive or evaluative meanings (but as Palmer already pointed out, meaning cannot be dissected into component parts); fourth, synonymous words may differ in their collocational distribution; and fifth, a set of words may be what Palmer calls synonyms in a loose sense: expressions whose meanings are close or overlapping, but not quite the same.

Lipka (1990: 63-67), again, defends the terms *denotation* and *connotation*, although he admits that the term *connotation* appears in various senses in literature. According to the writer, the term pair is useful, if *denotation* is used to refer to the denotative core meaning, and *connotation* to

any additional meaning properties of an expression: thus synonyms can be said to be marked by connotations (Lipka 1990: 64). Quite obviously, connotations can be analysed and classified if need be, as Lipka reminds (1990: 67-68). In addition to stylistic, affective and emotive overtones – which commonly are referred to as connotations – words may, for instance, be marked by temporal, regional and social connotations, following the dimensions of linguistic variation in general (Lipka 1990: 67-68).

However, quite similarly to Palmer, Lipka concedes that in some cases it is difficult to determine whether two words are synonyms with connotative differences or two denotatively distinct words, and cites a problematic case originally introduced by Leech (cited in Lipka 1990: 64, 20-21): are the words *horse*, *steed* and *nag* synonyms or not? Lipka is not able give an explicit answer, but concludes that the case illustrates the fuzziness of meaning (1990: 21). Here Lipka's thoughts coincide quite well with the claim made by Palmer that cognitive meaning and expressive/evaluative meaning are intertwined.

In this study, I will use the terms *denotation* and *connotation* for practical reasons. First, they are short, formally simple terms. Secondly, I find the ambiguity criticized by Palmer to be an advantage, as the term *connotation* can be used to cover any additional aspects of meaning that a word may have besides the denotative, core meaning. Similarly, the concept of synonymy is in this study to be understood in a broad sense as a clearly identifiable similarity of the denotative meaning.

2.3 Usage Labels in Dictionaries

Dictionaries, of course, describe features of linguistic variation at the lexical level. Usually this is done by means of usage labels, but as Landau (2001: 217) observes, special notes as well as qualifications within definitions may provide similar pieces of information. Some writers, Verkyul et al. (2003: 309) and Burkhanov (2003: 107) among others, emphasize the utility of illustrative

sentences as a supplement to labels. According to them, dictionary users are able to see how a given word is actually used and in what types of contexts it appears.

In this study, I will focus on usage information which is explicitly stated either in the form of labels, notes or qualifications within definitions; in other words, information which is openly stated and leaves little room for interpretation by the dictionary user. Thus, illustrative sentences are excluded from the study. Further, it should be stated once more that for the sake of convenience, the terms *label* and *usage label* are used in a general sense to cover information found in notes and definitions, as well.

Hartmann and James state that the purpose of usage labels is to show that a given expression is "typical or appropriate in a particular context or language variety" (1998: 150, s.v. *usage label*). Quite the contrary definition is also possible: usage labels are said to specify "restrictions" or "limitations" on the use of an expression (e.g. Hartmann and James 1998: 40, s.v. *diasystematic labelling*; Landau 2001: 215; Burkhanov 2003: 105).

In semantic terms, Hartmann and James (1998: 28, s.v. *connotation*) note that labels are used for indicating connotative aspects of word meaning, whereas definitions explain the more objective aspects of meaning: in other words, the denotation. Conversely, Verkuyl et al. (2003: 298) argue that usage labels should not be compared to the actual definition. In their view, a usage label is a meta-linguistic device that "restricts the definition to a certain context" (Verkuyl et al. 2003: 298).

Verkuyl et al. (2003: 298-299) point out that the practical significance of usage labels is evident above all when a dictionary is used for productive purposes: labels guide the dictionary reader in making right choices between expressions when speaking or writing, and warn the reader about the possible "social consequences" (Verkuyl et al. 2003: 299) of using a given expression.

Verkuyl et al. (2003: 306) arrange the most common types of usage labels in the following manner:

1. Group labels

    1.1. Geographical labels, for example: *American English*, *South-African English*, *regional*, *dialect*, *northern*, *African*, etc.

    1.2. Temporal labels

        1.2.1. First occurrence: *modern*, also date (sometimes it is also possible to state an exact date)

        1.2.2. Last occurrence: *obsolete*, *archaic*, also date (sometimes it is also possible to state an exact date)

    1.3. Frequency labels: *frequent*

    1.4. Field labels (professional or social domain): *aeronautics*, *alchemy*, *poetical*, *technical*, etc.

2. Register labels

    - Interdependent aspects of formality, offensiveness, figurativeness and aspects of mode of text: *colloquial*, *slang*, *jocular*, *derogative*, *vulgar*, *archaic*, *literary*, *euphemistic*, *figurative*, *pejorative*, *written*, *spoken*, *informal*, *formal*, *poetic*, *biblical*, *sexist*, *disparaging*, *coarse*, *rude*, *offensive*

    - Suggested system: *-2* (for *very informal*), *-1* (for *informal*), *0*, *1* (for *formal*), *2* (for *very formal*)

The arrangement consists of two main classes based on the user/use distinction also found in Quirk et al. (1985). Thus the first main class – group labels – contains labels that indicate if a word typically belongs to a particular group of language users, whereas labels of the second main class – register labels – direct the speaker in choosing between equivalents in different situations of language use. However, a closer comparison of Verkuyl's and his colleague's arrangement against Quirk et al. (1985) reveals quite considerable differences between the two. Whereas Quirk et al. (1985) name regional and social background as the main sources of variation which relate to the language user in particular, Verkuyl et al. (2003: 299-302) list region, time, frequency and subject field as user related types of variation. Further, their arrangement does not include factors of social background at all.

It is quite easy to understand that the time dimension belongs to group labels, if the relevant labels are interpreted as 'this word is or was used by people who are or were alive in the specified time'. Surprisingly, Verkuyl et al. (2003: 306) list the label *archaic* in both temporal labels and register labels. When left without an explanation, the reader remains puzzled. Svensén (1993: 178), however, gives one possible explanation: temporally marked expressions may have a stylistic function as they sometimes are used in order to give a solemn, high-style impression or alternatively to give a jocular effect. Moreover, placing frequency labels in the category of group labels strikes a little odd, too, and requires further explanation, which Verkuyl et al. do not give.

Yet another difference between Quirk et al. and Verkuyl et al. is their treatment of variation according to the subject field. Quirk et al. (1985: 16) claim that such variation is determined by the situation of language use, whereas Verkuyl et al. (2003: 301-302) consider field labels as a sub class of user labels. Of course, both views can be justified, depending on how special languages are defined. A definition according to Verkyul and his colleagues would probably hear: 'this word is used by specialists of the subject field'. Conversely, adopting the view of Quirk et al., the definition would be: 'this word is used when discussing matters belonging to the subject field'.

The second main class of usage labels in Verkuyl and his colleague's arrangement are register labels. It is a mixed class of labels that combine aspects of formality, offensiveness, figurativeness and mode of the text. The writers justify such a heterogeneous group by stating that subclasses would "create an inappropriate sense of exhaustiveness and independence" (2003: 306-307).

As an alternative to the complex set of traditional labels, Verkuyl et al. suggest a five step scale of labels: *-2* (for very informal expressions), *-1* (for informal), *0* (for common), *1* (for formal) and *2* (for very formal). The new system emphasizes that language varies into two directions from the neutral. Yet another feature in favour of the scalar labels is, according to the writers (2003: 303, 310), that the labels indicate how much and into which direction a particular word deviates from its neutral equivalent.

Another classification of usage labels is presented in Hartmann and James (1998: 150, 38-40). Their account of *diasystematic labelling* is a modification of the one given in Hausmann and Wiegand:

1. Currency or period (diachronic information): *archaic*, *obsolescent*, *new*, *in vogue*

2. Emotionality or attitude (diaconnotative/diaevaluative information): *appreciative*, *derogatory*, *offensive*

3. Frequency of occurrence (diafrequential information): *very frequent*, *becoming rare*, *rare*, *very rare*

4. Assimilation or contact: (diaintegrative information): *foreign*, *borrowed*, *assimilated, vernacular*

5. Mediality or channel (diamedial information): *written*, *spoken*

6. Normativity or standard (dianormative information): *incorrect*, *substandard*, *illiterate*

7. Formality or register (diaphasic information): *elevated*, *formal*, *informal*, *intimate*

8. Style or social status (diastratic information): *demotic*, *slang*, sometimes also *high*

9. Technicality or subject (diatechnical information): *law*, *music*, *chemistry*, *botany*

10. Textuality or genre (diatextual information): *poetic*, *conversational*

11. Regionality or dialect (diatopic information): *dialect*, *American English*, *British English*

This classification consists of eleven dimensions of usage, which are seen as scales where the opposite ends represent the extreme marked varieties and the neutral, unmarked variety is usually found in the centre (Hartmann and James 1998: 151, s.v. *usage labels*). For example, the scale for emotionality or attitude and the corresponding labels range from the utmost positive attitude (labelled *appreciative*) through the neutral to the other end of the utmost negative attitude (*derogatory*, *offensive*). Hartmann and James (1998: 39, s.v. *dianormative information*) suggest, however, that in some dimensions, the unmarked neutral is in fact situated at one end of the scale from which the marked varieties deviate only into one direction. This happens in the dimension of normativity, for instance. The correct, standard variety is unmarked, but deviations from this are found in one direction only; the labels indicating such a deviation are *substandard* and *illiterate*. The boundaries between some categories are not entirely clear-cut: Hartmann and James (1998: 40,

s.v. *diatextual information*) point out that the dimensions of formality, style and textuality are interrelated and overlapping.

The focus of this study lies on such usage labels that could be named as *cautionary labels*. They are labels that warn the dictionary user against offending the fellow communicator by using inconsiderate language and eventually giving a negative impression of the speaker. Verkuyl et al. list such cautionary labels in the group of *register labels*. In Hartmann and James (1998), the relevant labels are situated in the negative side of the dimension of *emotionality/attitude*.

Naturally, nearly all labels could be said to be cautionary. For example, temporal labels such as *archaic* or *obsolete* could be interpreted as 'do not use this word because it makes your speech sound unusually old-fashioned'. Similarly, the label *formal* could be said to imply 'do not use this word in casual conversation'. However, in this study the term *cautionary label* is used to cover labels that warn against offending the listener's feelings and thereby giving a bad impression of himself or herself. The label *euphemism*, although it is closely related to the concepts of taboo and controversial words, is regarded as an affirmative label meaning roughly 'use this word instead of another' and is therefore left outside the scope of this study.

3. Earlier Studies on Dictionary Labelling

The idea for the present study arose from Norri's article on labelling of derogatory words (2000), in which he pointed out a need for a chronological study on the issue. Consequently, the methods of this study are largely adopted from Norri's study. Similarities and differences between the two are detailed in Chapter 6 (Methods of Analysis).

Wachal's article (2002) gives one view of chronological changes in treatment of taboo words in dictionaries. The purpose of Wachal's study was to find out whether dictionaries have properly responded to changes that have taken place in the language as used in the mass media.

3.1 Derogatory Words in Dictionaries

The aim of Norri's study is to examine how different dictionary makers have tackled problems in marking negative connotations of words. The writer lists a number of "pitfalls" related to labelling. First he cites Landau's observation that politically influential groups are sometimes able to affect the labelling practices so that unpleasant words referring to them are labelled more rigorously than words referring to other, less influential groups of people (Norri 2000: 72). Another factor which complicates dictionary labelling, according to Landau, is that offensiveness depends on the context (cited in Norri 2000: 72-73); for example, a word that is normally considered offensive, may appear to have a positive overtone in some situations. Further, Norri adds that temporal and regional variation of language represent potential pitfalls in the labelling process, (2000: 73). In other words, meanings of words as well as their connotations change over time, which must be taken into consideration. Likewise may meanings and connotations differ depending on the geographical area where the word is being used.

Norri approached the problem area by investigating the use of attitudinal labels and usage notes in ten dictionaries. Seven of the dictionaries were British and three American. Two of the British dictionaries were targeted for language learners: *OALD4* and *Collins COBUILD English Dictionary* (abbreviated as *Cobuild*) (Norri 2000: 73). The purpose of selecting dictionaries that

originate from different parts of the world and that are intended for different audiences was to bring out possible variation in the labelling policies (Norri 2000: 75). Only the most recent edition of each title was investigated. The 145 words, whose entries were examined, were collected from a number of sources with the presumption that native speakers find the words offensive (Norri 2000: 74). In order to weigh the validity of Landau's argument about different groups being treated unequally in dictionaries, Norri collected words in seven semantic groups; these were nationality, racial or cultural group, people considered to be lacking in intelligence, deceitful people, sexual orientation, derogatory words for women, and derogatory words for men (2000: 74-75).

As to the results, Norri found out that dictionaries, indeed, differ in their labelling practices. The two learner dictionaries add a label considerably more often than the dictionaries aimed at a wider readership. *OALD4* attaches a label to 69% of the words it lists and *Cobuild* to 59%. For the other dictionaries, the figures range between 29–38%. Even though all but two compilations (*The Oxford English Dictionary* (*OED*) and *Webster's Third New International Dictionary of the English Language*) use a simple label – in other words, a one-word unqualified label – in the majority of cases, modifying adverbials and combinations of alternative labels are also found. One and the same word may be described as "often, sometimes or occasionally derogatory" in different works, which indicates that the choice of the frequency adverbial is not a simple task (Norri 2000: 91).

Norri (2000: 77) observed that many dictionary makers draw a distinction between *derogatory* or *disparaging*, on one hand, and *offensive* on the other. The first pair of labels is considered to emphasize the speaker's or writer's desire to insult, while the latter label describes the hearer's or reader's reaction to the use of the word. Thereby an expression may be both derogatory and offensive, but it is also possible to be offensive without the intention of being derogatory or disparaging. Norri (2000: 92) is of the opinion that the distinction does not hold in practice in the examined dictionaries. One and the same word may be labelled as *derogatory* in one work, *offensive* in another and both *offensive and derogatory* in a third one. Most often, such double labels

are found in the two learner dictionaries (*Cobuild* and *OALD4*) and in two American dictionaries (*Random House Compact Unabridged Dictionary* and *American Heritage Dictionary of the English Language*).

The calculations show that *OED* is the most prolific in the presenting alternative, non-negative functions to words. Sometimes the alternatives given in dictionaries are not entirely successful, as the pairs of labels do not necessarily exclude each other (Norri 2000:92). The point is clarified with the pairs "*dated* or *derogatory*" and "*informal* or *derogatory*" – the writer means that if a word is considered dated, it can still be derogatory. Similarly, if a word is used in informal style, it is not likely to lose its derogatory tone.

Norri confirmed that there are considerable differences in the cautionary labelling of words belonging to different semantic groups, as proposed by Landau. In the groups of nationality, race and culture, the likelihood of a cautionary label is at its highest (Norri 2000: 91); the writer believes that this is an indication of those words being considered very insulting. In the group of sexual orientation, the dictionaries are divided into two (Norri 2000: 91): in three large dictionaries (*Collins English Dictionary*, *The Oxford English Dictionary* and *Webster's Third New International Dictionary of the English Language*) warnings against the use of words of this type are rare, whereas the rest of the dictionaries use a label in over 60% of the entries. The use of cautionary labels is lowest in the groups of derogatory words for women and for men, and level of intelligence and deceitfulness (Norri 2000: 91), but the two learner dictionaries receive higher figures than the other dictionaries for these four groups as well. Norri believes (2000: 91) that the editors of general dictionaries trust that the definition is enough to reveal the negative overtone of the word, while the makers of learner dictionaries perhaps recognize a need to be more explicit. Yet another reason for the rare use of labels for words for unintelligent and deceitful persons is, according to Norri (2000: 91), that these words (such as *dimwit* and *conman*) do not refer to any specific well-defined class of people, as words such as *frog* and *darky* do.

Norri also investigated how unanimously the dictionaries add a label to individual words. The result was that forty-seven of the 145 words were labelled in a similar manner in all those dictionaries that listed the word (Norri 2000: 91). In other words, each of the forty-seven words either was assigned a label in all dictionaries covering the word, or none of the dictionaries assigned a label. For thirty-one words, one dictionary acted differently from the rest. However, there are differences between the semantic groups in this respect, too. Norri (2000:91) found out that the dictionaries agree more on the need for a label in the groups of nationality, race and culture, level of intelligence and deceitfulness, whereas their policies vary more for words for sexual orientation, arrogant or aggressive behaviour and physical appearance.

As a conclusion, Norri (2000: 92-93) remarks that differences between dictionaries are natural, but also reminds that a dictionary should be built harmoniously. Consequently, methods of providing pragmatic information call for more thorough planning. Dictionary-makers should also focus on explaining the labelling system to the readers, Norri adds (2000: 93).

3.2 Taboo Words in Dictionaries

Wachal (2002) studied how the altered taboo status of certain words can be seen in the dictionary treatment of those words. The writer claims (2002: 195), first, that words relating to the body and bodily functions are used in a wider scale than in the past, and conversely, that along with the political correctness movement, some terms for ethnic groups are condemned by the cultural community more heavily than before. In order to validate the claims, Wachal cites statistical information on the frequencies of offensive language in the mass media and, as an example of highly respected situations of speech, the United States' Congress.

The first group of taboo terms – that is, words relating to the body and its functions – were collected from a group of college students and from the mass media (Wachal 20002: 195). The second set of words – ethnic and racial epithets – were derived from a work by Mencken. This produced lists of forty and twenty-seven terms, respectively. The entries for these sixty-seven terms

were examined in twenty American and British dictionaries which covered virtually the whole twentieth century: from *Webster's International Dictionary of the English Language* (1906) to *The American Heritage Dictionary of the English Language* (2000) (Wachal 2002:195).

As to the set of words relating to body and its functions, Wachal (2002: 197) observed that the dictionaries use a wide range of labels to describe one and the same word: for instance, the word *fart* is labelled *vulgar* in one dictionary, *coarse slang* in another and even *taboo* in the third. Further, Wachal (2002: 197) criticizes that the degree of offensiveness of a word is not properly clarified: words such as *fart*, *asshole* and *cunt* were assigned quite a similar set of labels in the dictionaries, even though they usually are considered to differ in acceptability. When Wachal examined subsequent editions of certain dictionaries, he discovered that the labelling of individual words did not show any clear trend (2002:197) and concluded that dictionaries do not reflect the changed use of offensive language. Moreover, it was pointed out that too few offensive body words are listed in the dictionaries: fifteen of the forty words were listed in fewer than seven of the twenty dictionaries (Wachal 2002: 197).

Just as offensive body words, names for ethnic and racial groups were described using a wide range of labels. But unlike body words, ethnic terms were well covered in the dictionaries; only two of twenty-seven words were missing from the majority of the dictionaries (Wachal 2002: 198). Wachal (2002: 199) also noted that changes in the labelling of these terms reflect the change in the language use: ethnic terms are now labelled more negatively than before. However, different groups of people are not treated equally in the dictionaries. The heaviest disapproval is on offensive names referring to African Americans, Jews and Asians (Wachal 2002: 201), followed by ethnic groups of European origin, while the terms *Canuck*, *Limey* and *Redskin* receive the lightest cautionary marking, the last of which the writer finds quite surprising. Wachal also noted that capitalization of ethnic terms vary between the groups (2002: 201), which the writer assumes originating from old prejudices.

In the concluding remarks (2002: 201), Wachal recapitulates that dictionaries have not succeeded in describing the language as it is used as far as terms for body and bodily functions are concerned; special attention must be paid to covering the terms that are nowadays commonly used in the mass media. As to the group of ethnic offensive words, the dictionaries have fared better, but certain deep-rooted prejudiced habits should be corrected, as the writer urges.

4. Controversial Topics and Vocabulary in Terms of Taboo

As explained, the terms *controversial topic* and *controversial vocabulary* are in this thesis used to refer to items that are socially controversial: that is, topics that arouse strong, negative reactions in some people. People hold mixed, often unwavering, views about the stylistic status of the relating words. Whereas hearing a given word – not to mention using it – can be an extremely unpleasant experience for one person, another may remain quite untouched by the event even though the word does not belong to his or her active vocabulary. A third person may think that the word in question is entirely felicitous and trying to avoid it would be hypocritical. Quite often, but not necessarily always, the word *taboo* is heard in discussions about such problematic words and topics that they involve.

Handbooks of English usage often contain a section on taboo words. The purpose is to warn a reader about risky topics that many people in the English speaking cultures find disagreeable. They advise the reader against discussing such topics or suggest how they can be approached safely. Swan (1980: 589) states that taboos of the English language have three sources: the Christian religion, sex and elimination of body waste, all of which can be too "dangerous, holy, magic or shocking" (1980: 589) to be freely spoken about. Todd and Hancock's (1986: 458-459) list of present-day taboo areas in English includes religion, sex, bodily excretions, disease and death, social stratification, age and weight. Claire (1998: 1) introduces yet a few new taboo topics; she states that sexual body parts, sexual activities, bathroom functions, excrement, disgusting things such as vomit and nose picking as well as upsetting things such as cancer, death, and dying are taboos for Americans. Claire (1998: 1) further points out that religion, politics and racial issues are subjects on which people have discordant views and which may therefore lead into arguments. Therefore, Americans usually avoid discussions of these things too, the writer concludes (1998: 1).

Leach (1964) explains why certain things become tabooed. Taboos are, according to Leach, "apparently irrational prohibitions and inhibitions" of both certain behaviour and of certain

types of language (Leach 1964: 24). Verbal taboos, which Leach subsumes under the term *language of obscenity*, are of three types (1964: 28). The first group are the so-called dirty words: words that involve sex and excretion. The second group consists of blasphemies and profanities, and lastly, the third type of obscenity is *animal abuse*, a term by which Leach (1964) refers to language use "in which a human being is equated with an animal of another species" (p. 28). *Pig*, *fox* and *bitch* are commonly known examples of this type of language use.

Leach is of the opinion that taboos originate in ambiguity. Our perception of the physical as well as social environment is in its original state a continuum of phenomena (Leach 1964: 34-35):

Figure 2 The unbroken line is a schematic representation of continuity in nature. There are no gaps in the physical world (Leach 1964: 35).

Language serves as a means of creating order into the world: similar things are grouped together under the same name, and names are used to distinguish between different things. Leach argues (1964: 35) that as the language community strives for as clear-cut boundaries between different groups of things as possible (in other words, easily defined names), the more ambiguous sections of the continuum will be suppressed and therefore remain unnamed. The suppressed parts of the environment then arouse interest and anxiety, which in turn leads to taboo (Leach 1964: 37). Leach illustrates the idea as follows:
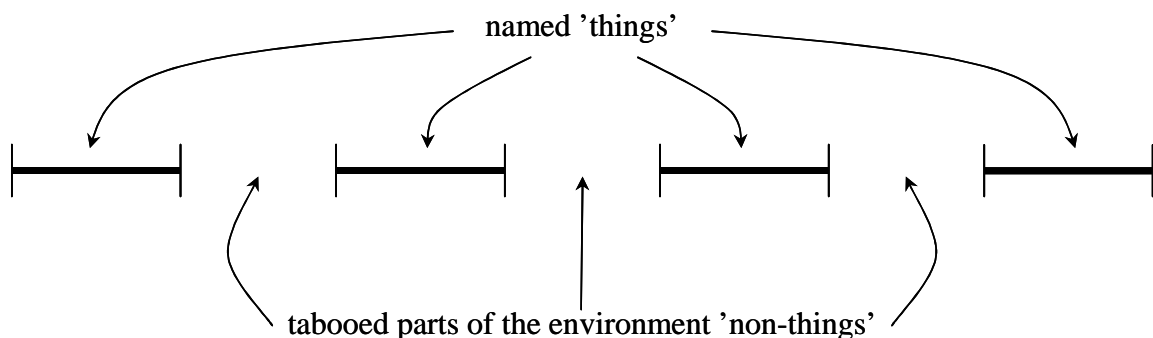
Figure 3 The relationship of tabooed objects to the world of names (Leach 1964: 35).

Further, Leach gives an example to support his theory. The exudations of the body – such as faeces, urine, semen, menstrual blood, hair and nail clippings – are commonly considered taboo. Leach argues that this is due to the ambiguous status of those substances being part of the self, on the one hand, and no longer part of the self, on the other. In other words, body exudations lie between the categories 'me' and 'the other world' (Leach 1964: 38). Yet another type of duality can be found concerning these substances. They are considered to be dirty, but also to hold magical medicinal properties. Leach (1964:38), however, notices that tears are not tabooed in the same way as the other substances. The reason for the different status of tears may be due to the fact that tears are not considered contaminating, Leach proposes (1964: 38 footnote).

Rothwell admits that Leach's taboo theory clarifies many of the seemingly irrational taboos, but criticizes it for failing to explain differences between items within a given semantic field (1982: 100-101). Rothwell sets forth the expressions *Lord!*, *sexual intercourse* and *kitty* against *God!*, *fuck* and *pussy*, and asks why the first set of words is generally regarded acceptable and the latter taboo. Similarly, the writer observes that people tend to judge sexual obscenities as more serious offences than excretory obscenities and profanities. Even though Rothwell (1982: 101-102) sees the truth in such claims, he emphasizes that definite comparisons of single taboo items are not valid, as considerations about what is taboo and offensive vary both in time and from person to person. Moreover, the context in which the taboo word is used – by whom, in what circumstances and to what ends – has an effect on how the word is taken (Rothwell 1982: 96 ff., 117).

Read (1934: 267) explains that there are different degrees of taboos within a semantic field, because some expressions are arbitrarily chosen as scapegoats to symbolize the forbidden, which then leaves their equivalents freely available to the speaker without fear of being affected by the power of the taboo.

Ullmann approaches the phenomenon of taboo from a different angle than Leach. The central idea of Ullmann's account (1962: 204-209) is that certain topics and words are avoided because bringing them up could have negative consequences. The origins of the term *taboo* are in a

Polynesian language where it referred to something forbidden; Ullmann (1962: 204-205) defines the term *taboo* as a ban on certain persons, animals and things as well as on their names, and proposes that these bans arise from three distinct emotions: fear, the feeling of delicacy and the feeling of propriety. Accordingly, verbal taboos can be classified into three groups. The first group consists of things which are feared: supernatural beings such as gods and evil spirits; animals which are believed to have supernatural qualities such as the bear, ants, bees and worms; and any other things that are associated with superstitious fears – for instance, the left hand (Ullmann 1962: 205-206). To the second group, to taboos of delicacy, belong topics that are considered unpleasant to talk about. Ullmann (1962: 206-207) cites illnesses, death, physical and mental defects as well as criminality as sources of this type of taboos. Finally, taboos of propriety are expressions that break the limits of decency and propriety. According to Ullmann (1962: 208), considerations of decency and propriety tend to change, but taboos of this category do generally involve sex, body parts and functions, and swearing.[2]

Allan and Burridge (1991) offer yet another explanation to taboo. The writers claim that whereas the term *taboo* originally referred to behaviour that was avoided for fear of punishment by a supernatural power, the term now usually applies to things that are considered distasteful rather than dangerous (Allan and Burridge 1991: 3, 12). Accordingly, Allan and Burridge (1991: 3, 12) suggest that taboos are avoided not for fear of punishment by a supernatural power but by fellow human beings; the punishment is losing one's face in they eyes of other people. The argument is based on Goffman's idea of *face-work*, which holds that social interaction normally aims to keep up a good image – *face* – of all parties involved (Allan and Burridge 1991: 5-6). Consequently, Allan and Burridge (1991) arrive at somewhat unconventional definitions of *euphemism* and *dysphemism*: "A euphemism is used as an alternative to a dispreferred expression, in order to avoid possible loss

---

[2] Norri (lecture 18 October 1999) has noticed that dictionary labelling of taboo words seems to vary according to Ullmann's taboo classes. He proposes that words involving taboos of fear do not usually receive a cautionary marking, contrary to taboos of propriety which are quite systematically equipped with a label such as *taboo* or *vulgar*. Taboos of delicacy are, again, often marked with a label (for instance, *offensive*) or another indication of a possible negative overtone, Norri observes.

of face: either one's own face or, through giving offence, that of the audience, or of some third party" (p. 11). A dysphemism is used when the purpose of communication is the opposite, hence the definition: "A dysphemism is an expression with connotations that are offensive either about the denotatum or to the audience, or both, and it is substituted for a neutral or euphemistic expression for just that reason" (p. 26).

Tabooed and distasteful topics in most English speaking cultures currently include, according to Allan and Burridge (1991: 12), bodily effluvia, reproductive processes and the corresponding body parts. However, the writers (1991: 12, 232) point out that taboo topics vary both in place and in time. For instance, whereas direct questions about another person's salary and personal income are regarded tactless in many parts of the world, in some cultures of the Far East such as Hong Kong, Singapore or Philippines one's income is said to be a normal topic of discussion (Allan and Burridge 1991: 12). As an example of temporal changes in considerations of taboo, Allan and Burridge remark that the old taboos of blasphemy, profanity and obscenity seem to have loosened, while "ageism, sexism, racism and religiousism" (1991: 232) gain ground as new taboos.

Allan and Burridge remind that dispreferred expressions are not always taboos; euphemisms are also used to make ordinary things sound better: for example, *a cozy home suitable for renovation* and *vermin control officer* sound more agreeable than *small dilapidated cottage* and *rat catcher* do, although the latter ones are not true taboo terms (Allan and Burridge 1991: 12-13).

There are certain parallels between Ullmann's and Allan and Burridge's accounts of taboo. This becomes apparent when taboos caused by fear of a supernatural power are set aside, and the remaining two of Ullmann's taboo classes (taboos of delicacy and taboos of propriety and decency) are brought into focus along with Allan and Burridge's motivation for avoiding dispreferred expressions (which is maintaining a positive image of others as well as of oneself). A few examples will perhaps clarify the point. Acting against the rules of propriety and decency – say, using rough language in company where polite behaviour is expected – poses a threat to the speaker's own

public image whereas the hearer's face remains unaffected. Further, a speaker who ignores the rules of delicacy can be said to pose a threat the public image of the hearer or of some third party. Nevertheless, in the end the speaker's own face, too, will be damaged as he or she is judged as tactless and ignorant of the rules of social conduct.

Yet another feature of taboo topics is their constant need for fresh euphemisms. Allan and Burridge (1991: 22) observe that as a euphemism becomes contaminated by its denotatum and loses its euphemistic value, it eventually turns into a taboo term. A new term will then emerge to perform the euphemistic function. Such a change in the connotation or emotive meaning of a word is commonly called *pejoration* or *deterioration* in works of linguistics (for instance, Gramley and Pätzold 1992: 33-34; Hughes 1988: 12).

5. Material

This chapter begins with a brief introduction to the history of the *OALD* dictionaries. I will also explain why this series of dictionaries is worth investigating. The second part of this chapter offers a detailed presentation of the sample material on which the analysis is made.

5.1 *The Oxford Advanced Learner's Dictionary*

The Oxford series of learner dictionaries, *OALD*, was quite a natural choice as material for studying diachronic changes in products of the branch of publishing; it has the longest history in the market of EFL dictionaries, measured both in actual time since the first edition appeared and in the number of editions published so far.

The history of *OALD* goes back to 1942 when A. S. Hornby as the general editor produced a dictionary called *Idiomatic and Syntactic English Dictionary*. In 1948, this work was republished by Oxford University Press under the title *A Learner's Dictionary of Current English*, the occasion of which is generally considered as the birth of the Oxford series of dictionaries for advanced foreign learners (e.g. Cowie 1998a: 251-252; Strevens and McArthur 1992: 593, s.v. *learner's dictionary*). To date, the dictionary has appeared in seven editions, although with some changes in the title of the work. The second edition, entitled *The Advanced Learner's Dictionary of Current English* (*OALD2*), appeared in 1963. From the third edition to the sixth one, the title remained the same, *Oxford Advanced Learner's Dictionary of Current English*. Those editions were published in the following years: *OALD3* in 1974, *OALD4* in 1989, *OALD5* in 1995 and *OALD6* 2000. The seventh edition came out in 2005 under a new, shortened title *Oxford Advanced Learner's Dictionary*. The latest edition is not taken into account in the present study, as the material was collected before its publication.

For quite a considerable time, the *OALD* series was "*the* English dictionary for advanced learners" (Rundell 1998: 318); it continued as the only brand in the market until 1978 when Longman brought out its learners' dictionary (*Longman Dictionary of Contemporary English*). That

and another work, which appeared in 1987 (*Collins COBUILD English Language Dictionary*), have so far been considered as the major challengers of *OALD*'s dominance (Strevens and McArthur 1992: 594, s.v. *learner's dictionary*), even though a number of other titles have also come available in the recent years. Against the seven editions of *OALD* published to date, *Longman* has been published in four editions and *Cobuild* in five editions.

Being the earliest product in the market, *OALD*'s standing is interesting: one the one hand, it is complimented on having established the standard of EFL lexicography for the other dictionaries to follow (Cowie 1998b: 249). On the other hand, the *OALD* series is criticized for being conservative (Allen 1996: 47) and for falling behind in user-friendliness in comparison to its challengers (Herbst 1996: 322). Although no comparison is made in this study to compilations by other publishers, it is interesting to see the impact of the landmark years of 1978 and 1987 on the Oxford series of learner dictionaries. Yet another landmark worth considering is the publication of *OALD5* in 1995, as *OALD5* was the first edition to be based on a corpus.

The six editions were put side by side to follow the evolution of the series through the years. Each edition of *OALD* has been issued in several impressions. All dictionaries examined in the present study were unrevised impressions except for *OALD3*, which was revised and reset in its eleventh impression in 1980. I used the seventeenth impression from the year 1983.

Information on policies of word treatment in the *OALD* series were sought for both in the introductory sections – prefaces and user guides – and where relevant also other extra material that the dictionary has to offer, but the most important source of data was a sample material collected from the dictionary entries. Prefaces and user guides, though apparently not very often read by actual dictionary users (Delbridge 1987: 67-68; Landau 2001: 148), usually provide information about the aims of the work, its intended users and describe the strands of vocabulary covered by the dictionary, as well as advise how to make use of the dictionary. Thus, they may give some answers to the research questions from the point of view of the dictionary-makers.

5.2 Method of Compilation of the Sample Material

The purpose of the sample material was to display the treatment of controversial topics and words in practice. The material consists of dictionary entries for 272 words which belong to eight semantic fields. There are words involving the human body and its functions, sex, names for ethnic minorities, physical and mental characteristics that are often found to be negative, as well as curse words. As the aim is to shed light on how entire semantic fields are represented in the dictionaries, not only the stigmatized members of those fields, the sample material also includes words that are fully accepted in the neutral language use. The topics were chosen to represent the diverse types of controversial topics, and thus bring out differences in the treatment of different taboos. The purpose was also to find words common enough for a foreign learner not only to encounter but also to use now and then when communicating in English. The eight meaning areas are:

1. words for a prostitute
2. words for a homosexual person
3. words for sexual intercourse
4. words for a person's buttocks
5. exclamations of anger
6. words for a person of an ethnic minority: an African American and a Native American person
7. words for an unintelligent person
8. attributes describing a fat person

Seven of the word sets were compiled by means of seven thesauri dating from the 1930s to the year 2000. One of the word sets, exclamations of anger, was put together using a strategy other than the thesauri. The reason for using several thesauri that extend over seven decades was to generate as many equivalent words as possible, including ones that are perhaps no longer current. A word that was considered common and neutral was chosen as a lookup word when searching for related words in the thesauri. For instance, when collecting words for the semantic field 'homosexual', the word was taken as a starting point. If that word did not seem to lead to a successful result, another lookup word (e.g. *gay)* was used for that work. To illustrate the point, Appendix B shows the thesauri entries that served as a basis for the lookup list for the semantic field 'homosexual'.

Although the goal was to form a well-balanced picture of each semantic field, it is unlikely that the lookup lists are completely exhaustive. It is the nature of thesauri to list words in clusters in which the items deviate to different extents from the central meaning. Judgements about where to draw the line between words to include and words to leave out are largely based on intuition, which increases the risk of relevant words being ignored. Similarly, some relevant words may not have appeared in the thesauri and have been ignored for that reason.

The opposite evidently took place, as well; the original lookup lists contained words that were later excluded from the study because the meaning given in the dictionaries deviated too much from the central meaning. It must be noted that, again, the decisions about where to draw the line between included and excluded meanings are, somewhat subjective. For example, in the semantic field 'homosexual', words such as *nancy* and *queen* were taken into consideration, whereas *punk* and *butch* were left out, as the words do not actually mean a homosexual person. Similarly, words which may very well indeed be used in the sense 'homosexual', but which did not appear in any of the Oxford learners' dictionaries, or did not appear in the dictionaries in that particular sense, were excluded from the final calculations. Among those are *woofter* and *lez*.

Occasional additions to the lookup list were made along the way of collecting data when related words turned up. Relevant pieces of information found in the dictionary entries were recorded in the tables found in Appendix A. The contents of the tables are explained in more detail in Chapter 6.

The method of compiling the word lists and excluding words from them has two consequences which makes this study different from the one conducted by Norri (2000). First, Norri compiled the word lists with the intention of collecting offensive words only. In this study, also inoffensive words are taken in. This has the consequence that the proportion of words with a negative label is likely to be lower than in Norri's study. Secondly, Norri did not make any exclusions from the word list based on the fact that none of the dictionaries did not list the word in question. In fact, his range of dictionaries was so wide and varied (ten dictionaries, originating both

in Britain and in America, targeted for different groups of audiences), that all of the 145 words studied were listed in at least one of the dictionaries.

Words used as exclamations of anger form a diverse group of items that seem to have only the expressive function in common. As thesauri list words by the meaning, and not by the function, the thesauri turned out not to be a handy source of exclamations of anger. Moreover, although many of the exclamations are semantically connected with the taboos of sex, body and religion, an examination of relevant thesauri articles shows that there usually are no subsections for interjections or exclamations in the same vein as there are subsections for nouns and verbs. Consequently, the set of exclamations of anger was compiled using an electronic format of *OALD6* that allowed searching for words in different ways. It was, for instance, possible to list words by the part of speech, or to sift out all entries containing a given word or a combination of words.

A preliminary investigation showed that definitions for exclamations in *OALD6* are usually quite similar: "used to express surprise or anger" (s.v. *blimey*) or "used as a swear word when sb is disagreeing with sth, or when they are angry about sth" (s.v. *bollocks*). In order to capture as many exclamations as possible, quite a complex search string was formulated with the help of Boolean operators: (use OR used) AND (anger OR angry OR annoyed OR annoyance). This yielded a list of 145 items. Subsequently, thirty-four items which, according to the entry information, can stand as independent one-word utterances were selected to the sample material. Thus, longer idioms such as *"For Christ's sake!"*, *"Can you beat that/it!"* or *"For crying out loud!"* were omitted. Similarly, adjectives and adverbials such as *blinking*, *bloody* and *blooming* were left out. As an exception to the one-word rule, a number of verbs that are often used transitively in curses are taken into consideration (as is *"Damn it!"*), as they also appear as one word exclamations (*"Damn!"*). The intention was to augment the list with any exclamations cropping up in cross references and such in the process of collecting the data. However, only one such item was found (*flip* in *OALD5*). Thus, the sample material for exclamations of anger consists of thirty-five items altogether.

Having *OALD6* as the main source of the word list has some likely effects on the results. First, the present-day exclamations are overrepresented in comparison to exclamations that were current in the past decades. In comparison to the other editions, *OALD6* is likely to fare better as to the number of words than it really deserves. Second, items that have been used in the past but are no longer current, and therefore do not appear in *OALD6*, are largely ignored, even though they may appear in the older editions. Consequently, the older editions do not probably reach as good results as they deserve. Third, the distortion increases the earlier the edition is.

6. Methods of Analysis

The analysis of the sample material seeks to answer the following research questions:

1. How has the inclusion of words of controversial topics changed from *OALD1* to *OALD6*?

    a. How many words are listed in each dictionary?

    b. How consistently are the words kept listed?

2. How has the labelling of words of controversial topics changed from *OALD1* to *OALD6*?

    a. What labels are used in the dictionaries?

    b. How frequently do the dictionaries use labels?

    c. How frequently do the dictionaries use the taboo sign as a label?

    d. Do the dictionaries prefer unmodified or modified labels?

    e. Do the dictionaries indicate alternative, inoffensive, contexts of use?

    f. Do the dictionaries use multiple cautionary labels?

    g. How consistently do the dictionaries assign a label to individual words?

    h. How consistently do the dictionaries assign the taboo sign to labelled words?

The method of analysis is based on the one used by Norri (2000) in his study on labelling of derogatory words, but a number of alterations have been made and some additional calculations have been adopted to serve the purposes of this study. For example, the special taboo sign used by the *OALD* series (marked with △! in the present work) receives special attention not found in Norri's work. All the relevant information for the analysis is found in the tables of Appendix A. What follows is an explanation of the contents of the tables.

Appendix A contains three tables for each of the eight semantic fields (Table 1a - Table 8c) and seven tables that combine and contrast information from the eight categories (Table 9a - Table 9g). The first table for each semantic field (Table 1a, Table 2a and so on) summarizes the relevant information gathered from the dictionary entries for each word under investigation. The shorthand representation of the entry contents follows the one employed by Norri (2000):

Semantic field	*Contents of a dictionary entry*	*Marking in the tables of Appendix A*
homosexual (as a noun)	**queer** *adj.* ❶ strange; odd, as *a queer way of talking*. ❷ open to suspicion; doubtful; shady, as *a queer character*. ❸ out of sorts; unwell; faint, as *to feel queer*.❹ eccentric; slightly wrong in the head. **(find oneself) in Queer Street**, in debt; in trouble. –*vt.* (colloq.) spoil; put out of order. **queer a person's pitch**, see ²*pitch*. **queer-ly**, *adv.* **queer-ness**, *n*. **queer-ish** *adj.* somewhat queer. (*OALD1*)	0
homosexual	**queer** *adj.* **4.** (modern use, of a man) homosexual. […] *n*. homosexual. (*OALD2*)	–
homosexual	**queer** *adj* […] *n* (*sl derog*) homosexual man. (*OALD4*)	+
homosexual	**queer** *adj* […] *n* ⚠ (derog sl) homosexual. (*OALD3*)	+
homosexual	**queer** *noun* (⚠**,** *slang*) an offensive word for a HOMOSEXUAL, especially a man, which is, however, also used by some homosexuals about themselves. (*OALD6*)	(+)
homosexual	**dike** (also **dyke**) *n* (*sl usu offensive*) a LESBIAN. (*OALD5*)	(+)
prostitute	**harlot** *n* (*arch or derog*) prostitute. (*OALD4*)	(+)

Table 1 Conversion of dictionary information to a shorthand marking in the tables of Appendix A.

Table 1 above shows that *OALD1* does not list the word *queer* in the relevant nominal sense 'homosexual'. Therefore, the word is assigned a zero in Table 2a of Appendix A. Conversely, *OALD2* lists the nominal use of *queer*, but as there is no cautionary label attached, the word is marked with a minus sign. A plus sign means that the word is listed in the relevant sense and carries a cautionary label. As can be seen in the entries for *queer* in *OALD4* and *OALD3*, no distinction is made between entries that contain a single cautionary label (such as *derog*) and a multiple cautionary label (⚠ *derog*). The last three rows of the table above present cases in which the word

is marked with a bracketed plus sign. Such a marking stands for entries that, on the one hand, recognize a need for caution but, on the other hand, suggest that there are situations in which the cautionary label does not apply. For instance, even though *OALD6* both assigns the taboo sign to the word *queer* and describes the word as offensive, it also implies that the word is used by homosexuals about themselves without the disparaging tone. Similarly, *OALD5* indicates that the word *dike* is usually considered offensive. In other words, the dictionary suggests that the word may in some situations be used without offence, but does not specify what the situations may be like. The entry for *harlot* in *OALD4* specifies a choice between two uses, as it is labelled *arch or derog*. First, the word can be understood to be an old-fashioned word; thus, in old or old-fashioned texts the word *harlot* may appear to be attitudinally neutral. The other alternative, then, applies to current language use – the word is now regarded derogatory. The exact wording of the cautionary labels and any other information considered relevant is found in footnotes under each table of Appendix A.

The second table for each semantic field (Table 1b, Table 2b and so on) summarizes the contents of the previous table in a numerical format. The table shows the number of words listed in each dictionary, the number of words marked with a plus sign and the number of words marked with a bracketed plus sign. The categories of + and (+) are shown separately in order to highlight the preference for the simple type of label or for a modified label in each dictionary. The bottom row of the table shows the frequencies of labelled words as percentages of those words that are listed in each dictionary. Turning the absolute numbers of labelled words into percentages allows a comparison between the subsequent editions.

The third table for each semantic field (Table 1c, Table 2c and so on) focuses on the use of the taboo sign ⚠. The number of words labelled with the taboo sign as a single label and as combined with another cautionary label are shown separately. Finally, calculations are made about the percentage of words assigned the taboo sign as against all labelled words.

Table 9a and forward recapitulate information gathered in the eleven word groups so that the groups and the dictionaries may be compared against each other. The contents of these tables are discussed in the concluding chapters of the study.

Table 9a and Table 9b focus on the number of controversial words in dictionaries: these tables do not have equivalents in Norri's study (2000). Table 9a combines the data about word inclusion in each semantic field and each dictionary. In order to be able to compare the categories against each other, the actual numbers of words listed are turned into percentages of the total of words under investigation. The bottom row presents the overall percentages of word inclusion in each dictionary.

Table 9b deals with the consistency in which individual words remain listed in the dictionaries once they have been taken in the dictionary. The words are classified into four groups, plus a residue: 1) words that appear in all six dictionaries, 2) words that remain listed once they have made their way in, 3) words that are listed in *OALD1* but seem to have disappeared permanently at some point of time (in practice, the last time the word was listed was in *OALD4* or earlier) and 4) words that are regarded as fluctuating between being listed and left out again. The residue consists of words that made their first appearance in *OALD6* or were left out for the first time in *OALD6*. These words cannot be placed into the other four groups, as it remains yet to be seen whether they represent more stable changes or are instances of temporary fluctuations.

The next three tables (Table 9c, Table 9d and Table 9e) are adopted from Norri (2000), the last one being a modified version of the original. Table 9c recapitulates the frequencies of cautionary labelling in each semantic field and each dictionary, merely for the sake of comparison between the categories. Table 9d puts all semantic fields together and shows the overall frequency of labelling in each dictionary. Table 9e directs attention to the consistency in which individual words receive a cautionary labelling across the line of dictionaries. The words are classified into three groups – uniformity, near-uniformity and non-uniformity. A fourth group consists of a residue

of words that are listed in one dictionary only; in those cases, a comparison naturally is not possible.[3]

Uniformity in labelling means that either all dictionaries listing the word assign a cautionary label to the word or all dictionaries do not assign a label. Near-uniformity refers to those cases in which one dictionary acts in dissonance with the others. Non-uniformity occurs when there is even more disagreement over the need for a label between the dictionaries.[4]

Table 9f and Table 9g show summarize the information about the overall frequency and consistency of the taboo sign in particular.

---

[3] I have followed a course of action that differs somewhat from the policy of Norri's study (2000). Norri does not have a residue group, but included words of one occurrence into the group of uniform labelling.

[4] Another difference to Norri (2000) is found in the interpretation of uniformity of words with two occurrences, when one of the dictionaries uses a label and the other does not. Norri regards the labelling in such cases as being in near-uniformity, while I have grouped such words in the group of non-uniformity.

7. Policies of Word Treatment as Stated in the User Guide Sections of *OALD*

Dictionaries usually contain a number of pages in front of the A-Z section in which the editors explain the principles according to which the dictionary was compiled, and advise the reader of how to make the best use of the dictionary.

7.1 Policies of Word Inclusion

A. S. Hornby, the general editor of *OALD1*, states in his *Introduction* (p. iv) that the dictionary is designed to be of service to foreign students of English who have not yet advanced to university level in their studies. The dictionary is said to concentrate on words that a foreign learner is likely to meet. Most archaic words, as well as purely scientific and technical words, have therefore been left out. The writer assumes, however, that foreign learners will encounter certain colloquial and slang words in the texts they read. A number of words belonging to these areas of vocabulary, as well as common foreign and Latin expressions, have therefore been included in the dictionary. These groups of words may, of course, contain socially controversial words, but this aspect of vocabulary is not taken up in the prefatory section. Nor does the writer specify how words belonging to the more peripheral areas of the vocabulary – colloquial or slang words, for instance – are distinguished from the common core of the vocabulary.

The writer of the *Preface* of *OALD2* (p. v) tells that since the publication of *OALD1*, two smaller dictionaries have come out. As both of these were targeted on learners at more elementary levels, this new edition of *OALD* attempts to meet the needs of advanced students as well as teachers of English. As to the spheres of vocabulary that the dictionary covers, *OALD2* follows in the footsteps of *OALD1*. The writer states that the dictionary describes English as it is currently used (p. v). The number of entries has been increased, because a large number of new words and phrases have come into use since the first edition. In addition to the most central vocabulary, the dictionary registers a number of archaic, foreign, scientific and technical as well as colloquial and

slang words, if they are commonly found in texts read by foreign learners (p. v-vi). Similarly to *OALD1*, *OALD2* does not make any reference to the inclusion of socially controversial words.

In the user guide of *OALD3*, the intended user is defined as a learner who is able to read and understand texts of "moderate difficulty" (p. xiii). The guidelines of word inclusion are in line with the previous editions: the dictionary is said to list words that a learner is likely to meet in everyday language use. According to the writer (p. xiii), the dictionary records both written and spoken language, found in official as well as informal situations, and in contemporary literature. As in the previous editions, nothing in particular is said about the inclusion of socially controversial words in the dictionary.

OALD4 stands apart from the other editions with its two separate user guides: *A Practical Guide* for the benefit of student users in the front of the book and *A Detailed Guide to the Entries* at the back of the book aimed at the more advanced students and teachers of English. The *Practical Guide* gives the reader basic skills for finding relevant pieces of information and for interpreting and making use of the information correctly (pp. xiii-xix). The *Detailed Guide* (pp. 1545-1579), as the name implies, offers a more thorough exposition of the types of information found in the dictionary. It also takes up a number of lexical problems and illustrates how the dictionary may be of help.

Another difference to the previous editions is that *OALD4* does not make any reference to the principles of word inclusion. Neither of the two user guides informs about the types of words that are believed to be important for a language learner to know and consequently for a dictionary to account for.

Until the appearance of *OALD4*, each edition has expanded in the length and contents of their user guide sections. However, *OALD5* seems to take a new the direction. The user guide of *OALD5* is considerably shorter than the *Practical Guide* of *OALD4*, the more concise one of the two user guides of the previous edition. The chief editor Jonathan Crowther states in the *Preface* of *OALD5* (p. vi) that for the first time, the dictionary is built on the British National Corpus, a

computer databank of modern written and spoken language. In fact, nothing else is told about the types of vocabulary described in the dictionary.

OALD6 has a rather short user guide, too. Neither the areas of vocabulary nor the typical dictionary user are specified, but the chief editor Sally Wehmeier points out in the *Preface* (p. vi) that the learner-centred focus of dictionary making already adopted in the first edition is still sharpened in *OALD6*.

7.2 Policies of Usage Labelling

The introductory part of *OALD1* does not discuss the policies of usage marking. There is a general list of abbreviations used in the dictionary (p. xxv), among which some items obviously function as usage labels. For example, the list contains the abbreviation *emot.* which stands for '(used to mark words which rouse) emotion'. Usage labelling will not, however, be discussed further here; first, because the list evidently does not cover all labels used in the dictionary, and conversely, because the list of abbreviations contains items which are not usage labels. Thus, the discussion of usage labelling and other means of usage notification employed by *OALD1* will be based solely on the sample material as examined in Chapter 8.

As to the need and purpose of usage marking, *OALD2*, *OALD3* and *OALD4* state their views, whereas *OALD5* and *OALD6* do not address the issue. According to *OALD2* (p. xi), foreign learners may have difficulties in recognizing stylistic values of words. Nevertheless, the editors remark that giving explicit usage indications in dictionaries is "almost impossible" (p. xi), because there are many borderline cases in which it is difficult to determine which label is the appropriate one. The writer concludes that usage indications are given "only when stylistic values can be stated without much doubt" (p. xi). The writer of the *OALD3* user guide reminds that knowing the spelling, pronunciation and grammatical features of a word does not add up to being able to use the word correctly (p. xxvi); familiarity with factors of context and style is also needed. Therefore, the dictionary is said to provide information on any contextual and stylistic restrictions that words may

have (p. xxvi). The *Practical Guide* of *OALD4* addresses the issue of usage labelling only briefly. By means of an example, it explains how the dictionary helps the reader to make out differences between stylistically different variants (p. xviii). The *Detailed Guide* of *OALD4*, unlike the *Practical Guide*, explicitly recognizes the need for usage notification. The writer asserts that it is often as problematic for a foreign learner to use words appropriately as it is to know the exact meanings (p. 1572). To make the dictionary user aware of stylistic differences between words, the dictionary is said to use labels.

A list of labels used in the dictionary has been offered in every edition since *OALD3*. As already mentioned above, *OALD1* does not contain any information of labelling. *OALD2* does not produce a list of all labels, but it takes up a number of labels in its discussion on stylistic values of words. For example, the writer of the user guide informs that slang expressions are marked, because foreign learners are advised against using slang (p. xi). Similarly, the labels *taboo* and ⚠ (for vulgar or incorrect usage), are introduced as cautionary signs (p. xi). A foreign learner should never use words marked with these labels, even though they may be met in texts read by learners, the writer concludes. Using the cautionary sign ⚠ in another sense than 'taboo' (namely 'vulgar or incorrect') makes *OALD2* different from its successors. Nevertheless, in this study both the cautionary sign ⚠ and the verbal label *taboo* of *OALD2* are treated equally with the cautionary sign ⚠ used in the newer editions, because both stand for a strong opposition to the use of the word.

OALD3's inventory of usage labels consists of 58 labels for specialist English registers and 27 labels for stylistic values (See Appendix C for the lists of usage labels of *OALD3 – OALD6*). The labels of *OALD3* are not defined nor is their use explained in detail, but a few cases are taken up to show how labels can be of help in choosing the more usual or appropriate expression (p. xxvi); the meaning and use of a number of labels become clear by reading in the text. The taboo sign ⚠ receives more attention. According to the writer (p. xxvii), taboo words are used for their shock value. As such words may make the listener uncomfortable or upset, the writer discourages foreign learner from using taboo words.

The *Practical Guide* of *OALD4* (p. viii-xix) resembles the user guide of its predecessor: the use of labels is explained with a few illustrative cases only. Conversely, the *Detailed Guide* of *OALD4* (p.1572-1575) offers a comprehensive description of the labelling system. The labels are classified into six major groups and the label *propr* (for proprietary names) forms a minor group of its own (See Appendix C). The meaning of each label is interpreted, and a number of words are given as examples of words described by the label, which perhaps clarifies the meaning of the label further. In addition, typographical conventions as well as principles of combining and modifying the labels are specified in the *Detailed Guide*.

OALD5 and *OALD6* are quite alike in their introductions to labelling. Neither of the dictionaries goes further in their accounts than listing all labels used in the work (See Appendix C), defining them briefly and giving a few example words on each label. Typographical and combinatorial details or possible restrictions are not accounted for in either of the dictionaries. In *OALD5*, the labels are divided into three main groups (attitudinal and contextual labels; other restrictive labels; field labels). *OALD6* does not mention field labels; its two label groups are attitudinal or situational labels and other restrictive labels. *OALD6* deviates from the older editions in that it has abandoned abbreviated labels except for the taboo and trademark signs as well as the regional labels *AmE* and *BrE*.

An examination of the limited selection of labels used in the newest four dictionaries shows that cautionary labels are used in all, even though the title for these labels varies. In *OALD3*, the relevant labels are found in the group of labels for stylistic values; in *OALD4,* they are found in the groups of register labels and evaluation labels; in *OALD5* in the group of labels for attitude and context and in *OALD6* in labels for attitude and situation. The relevant cautionary labels are:

OALD3: *derog* (for derogatory), *emot* (for emotive), *pej* (for pejorative), *vulg* (for vulgar), ⚠

OALD4: *derog* (for derogatory), *offensive*, *sexist*, ⚠

OALD5: *derog* (for derogatory), *offensive*, *sexist*, ⚠

OALD6: *disapproving*, *offensive*, ⚠

The number of different cautionary labels has diminished from five of *OALD3* to three of *OALD6*. It is quite surprising that the taboo sign is the only label to appear in all four editions. Changes have been made to the other labels, even though it seems that certain basic functions are fulfilled by every set of labels. In other words, labels seem to have been replaced by others with a similar meaning. A closer look at the definitions given to the labels lends evidence to this. Definitions to the labels used in *OALD3* are imported from the dictionary entries, as the user guide does not define the labels as is done in *OALD4 – OALD6*:

⚠ *taboo* 'Some words in the Dictionary are followed by the sign ⚠. These are *taboo* words. They are words used when the speaker wishes to swear, or be indecent, or be offensive. They are all words that are likely to cause embarrassment or anger if they are used in the wrong situation. The learner of English is advised to avoid using them.'(*OALD3*, p. xxvii)

'[⚠] denotes words or senses likely to be thought offensive or shocking or indecent (though not necessarily by everyone or on every occasion), eg *wop*, *nigger*, *Christ!*; *fuck*, *prick*, *shit*, *piss*. Foreign learners should exercise great care in using these words. They should also note that words such as *wop* and *nigger* are generally used with the deliberate aim of giving offence.' (*OALD4*)

'Taboo words are likely to be thought obscene or shocking by many people and should be avoided by learners. Examples are *arse*, *bloody*, *shit*.' (*OALD5*)

'Taboo words are likely to be thought by many people to be obscene or shocking and you should avoid them. Examples are *bloody* and *shit*.' (*OALD6*)

derogatory 'tending to damage or take away from (one's credit, etc); insulting' (*OALD3*, s.v. *derogatory*)

''Derogatory' words, etc imply that one disapproves of or scorns the person or thing referred to or described by those words, eg *puerile*, *skulk*, *suspect* (adjective).' (*OALD4*)

'Derogatory expressions show that the user feels disapproval or scorn, eg *brat*, *fuddy-duddy*, *pedantic*.' (*OALD5*)

disapproving	'Disapproving expressions show that you feel disapproval or contempt, for example *blinkered*, *faceless*, *jumped-up*.' (*OALD6*)
pejorative	'depreciatory; disparaging; deteriorating in use or meaning' (*OALD3*, s.v. *pejorative*)
offensive	'This label denotes words used to address or refer to people, usually with a deliberate intention of offending them, especially on account of their race or religion. Words such as *spick*, *wop*, *nigger* are almost always used offensively in this way; words such as *arsehole* and *prick* are often found shocking, but they need not be used as terms of abuse.' (*OALD4*)
	'Such expressions are used to address or refer to people in a way that is very insulting, especially in connection with their race or religion, eg *nigger*, *wop*, *yid*.' (*OALD5*)
	'Offensive 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 *mulatto*, *slut*, *cretin*. They should be avoided.' (*OALD6*)
sexist	'This label denotes words and phrases that express a (sometimes unconscious) discriminatory or patronizing attitude towards someone of the opposite sex. They are almost always words, etc used by men about or to women, and can be used to express approval in a 'man-to-man' context, eg *dolly*, *dumb blonde*, *a bit of skirt/crumpet/all right*, *an easy lay*.' (*OALD4*)
	'Such words express an unfair or patronizing attitude towards a person of the opposite sex. They are usually used by men about women. Examples are *career girl*, *dolly-bird*, *looker*.' (*OALD5*)
vulgar	'ill mannered; in bad taste' (*OALD3*, s.v. *vulgar*)
emotive	'of, tending to excite, the emotions' (*OALD3*, s.v. *emotive*)

The views on taboo language of the four dictionaries vary somewhat. In the user guide of *OALD3*, the speaker's intention of being indecent or offensive is emphasized, whereas in the three newer editions, feelings of indecency and offensiveness are considered reactions of the listener. *OALD4*'s stand on the use of taboo words is milder than that of the other editions, as it urges the speaker to "exercise great care", whereas the others quite directly advise against using taboo words.

The labels *derogatory*, *disapproving* and *pejorative* seem to be related. *OALD3* uses both *derogatory* and *pejorative*, while the others only one of them. *OALD3* does not appear very user-friendly in this respect. As already noted, the user guide does not explain the meanings of the labels, nor does the difference between *derogatory* and *pejorative* clear up by looking up the relevant entry

articles. The situation is made more complicated by the fact the definition of *pejorative* contains many difficult words (that is, *depreciatory*, *disparaging*, *deteriorating*) which a foreign learner may not understand it without looking up also those words.

It is quite evident, however, that the label *derogative* of *OALD4* and *OALD5* has been replaced with *disapproving* in *OALD6*; the definitions of these labels contain essentially the same information. This may be an attempt to make the dictionary more user-friendly by a label with a more transparent meaning.

The label *offensive* is used in *OALD4*, *OALD5* and *OALD6*. *OALD4* and *OALD5* use also *sexist*, which appears to have quite a similar function as *offensive*. All three dictionaries tell that offensive words have people as referents. *OALD4* draws attention to the speaker's intention of offending, whereas *OALD5* and *OALD6* state that using offensive words is considered very insulting, thereby emphasizing the role of the listener perhaps. All three dictionaries agree on that race and religion are topics from which offensive words originate, but *OALD6* adds sex and disabilities to the list.

It appears that the label *offensive* of *OALD6* fulfils also the function that the label *sexist* has in *OALD4* and *OALD5*. The two dictionaries inform that the label *sexist* marks words that express an unfair, discriminatory or patronizing attitude toward persons of the opposite sex, usually men's attitude to women. *OALD4* remarks further that the speaker may sometimes remain unconscious of his or her negative attitude, and that sexist words are sometimes used in a positive, appreciative tone in discussions between men. To conclude, *OALD6* seems to treat sexist words as one subtype of offensive words.

The distinction between *derogatory* and *offensive* (see Chapter 3.1 on page 15) is not emphasized in the definitions of the labels in the *OALD* series. The explanation of *offensive* in *OALD4* works, in fact, quite against the claim, as it says that offensive words are usually used "with a deliberate intention of offending". Instead, the difference between *offensive* and *derogatory* (or *disapproving* in *OALD6*) that appears to be present in all four editions is the range of referents for

words labelled as offensive: they are said to refer to persons of certain groups such as religious groups, racial groups, either of the sexes or disabled people according to that particular group membership. The labels *derogatory* and *disapproving*, on the other hand, seem to relate to negative attitudes toward the referent based on any individual characteristic.

The two remaining labels of *OALD3* which are considered cautionary, *emot* and *vulg* (for emotive and vulgar), do not have parallels in the newer editions. *Vulgar* is regarded a cautionary label, because it can be understood to warn the speaker against giving a bad, ill-mannered, impression of himself or herself. The status of *emot* as a cautionary label is even somewhat uncertain, as the definition does not indicate explicitly what types of emotion do words marked with the label excite. The label is perhaps intended to cover both positive and negative emotions.

To conclude, there is a trend toward not only a systematic description of language, but also toward simplicity and user-friendliness. Fixed sets of usage labels, as well as definitions and examples on how labels are used are pieces of evidence of such a goal. Moreover, labels are no longer abbreviated or otherwise opaque in meaning.

8. Treatment of Controversial Topics and Vocabulary in *OALD*

Although the analysis of the sample material includes quantitative measures, it should be kept in mind that the numbers of words studied in each semantic field are relatively low. Therefore, definite conclusions based on the numerical data should be drawn with caution. For example, the semantic field 'prostitute' consists of twenty items of which *OALD2* lists eleven only. This means that each word equals to nine percentage points. *OALD4*, on the other hand, lists sixteen words, which equals six percentage points for each word. Considering these shortcomings, the percentages must be compared on a rough scale, showing trends of development only.

8.1 Treatment of Words for a Prostitute

Twenty words for a prostitute were included in the study. Prostitution has traditionally been considered an occupation for women only. Accordingly, words in this semantic field are most often used to refer to women, except when the opposite is explicitly indicated (e.g. *male prostitute*). In some cases, the form of a word reveals a female referent (e.g. *call-girl* and *streetgirl*). Another group of words have a gender-neutral form, but are nevertheless understood to denote women. For instance, the word *scrubber* is defined as "an offensive word for a prostitute or for a woman who has sex with a lot of men" in *OALD6*.

Definitions of some of the words in this category have changed, however – perhaps as a reaction to growing awareness of the existence of male prostitutes. The word *prostitute* itself serves as an example. Its definitions in *OALD1* and *OALD2* contain the word *woman*, which is then substituted for *person* in the newer editions. At the same time, the *OALD6* definition of *hooker* may be evidence of the fact that prostitutes being women is still considered the norm, as the definition reads: "a prostitute (= a woman who has sex with men for money)." According to the dictionary

definitions, two of the twenty words are used only of male prostitutes or paid lovers (*gigolo* and *rent boy*).[5]

Words for a prostitute are controversial for many reasons. Following Ullmann's classification of taboos (see Chapter 4, page 22), words in this semantic field involve both taboos of propriety and of delicacy: it is a question not only of sex, but also sex outside a steady relationship, paid sex, sometimes also illegal sex, and of behaviour often considered quite pathetic and morally reprehensible. A loss of face is an instant punishment for those who have become publicly connected with prostitution – the topic certainly meets the characteristics of a controversial topic and taboo.

8.1.1 Inclusion

A full account of the treatment of words for a prostitute is found in Table 1a, whereas Table 1b offers a summary of the previous table. The number of words listed in each dictionary in this word category ranges from eight to sixteen of the twenty words investigated. The lowest number of words for a prostitute appears in the oldest dictionary, *OALD1*, from which the number rises steadily until reaching the peak of sixteen words in *OALD4*. In *OALD5*, the number falls by four words, which is quite a significant change regarding the size of the word group. In the newest edition, *OALD6*, the number of words is increased by one, again.

When the inclusion of the words is considered individually, the situation appears somewhat more complex. Five of the twenty words are listed in all six editions (*courtesan*, *harlot*, *prostitute*, *streetwalker* and *whore*), and another five words remain listed once having made their way into the dictionary (*call-girl*, *gigolo*, *hooker*, *hustler* and *tart*). This suggests that these words are well established in this semantic field. The only word regarded as having disappeared is *white*

---

[5] The definitions given in dictionaries occasionally are rather inexplicit. For instance, *gigolo* is defined as 'hired dancing-partner or companion' (not a verbatim quote) in *OALD1-3*. Owing to the vagueness, it is sometimes difficult to determine whether a particular meaning is listed or not. I have followed rather a strict course in such cases, marking definitions which do not explicitly include the meaning under examination as '0'. As to this category of words, definitions which convey the meanings 'prostitute' or 'paid lover' are considered acceptable.

slave, as it has not been listed since *OALD4*. On the inclusion six of the words (*moll*, *scarlet woman*, *scrubber*, *streetgirl*, *trollop* and *woman of easy virtue*), the dictionaries vary more; the status of being 'listed' and 'not listed' in the relevant meaning changes more than once. The word *moll* may serve as an example of such fluctuation. The word does not appear in *OALD1* at all, receiving the mark '0'. In *OALD2* and *OALD3*, *moll* is defined as 'woman companion of a gangster; prostitute', clearly worth the mark '–'. In *OALD4* and the newer editions, the definition reads 'a female friend of a criminal' without any indication of the sense 'prostitute'. Therefore the mark is again '0'. The residue of this semantic field consists of three words: *rent boy*, *working girl* and *strumpet*. The first two words appear in the relevant sense for the first time in *OALD6*, whereas the same edition dropped the last one.

8.1.2 Labelling

The terms used as labels of a negative connotation in this semantic field are the following: *term of abuse* and *contemptuous term* (*OALD2*); *term of abuse*, *derog* and the taboo sign ⚠ (*OALD3*); *derog* (*OALD4* and *OALD5*); and *disapproving* (*OALD6*). In addition, *OALD6* integrates the cautionary note into the definition proper once in this word category, explaining the meaning of the word *scrubber* as "an offensive word for a prostitute". *OALD1* does not attach a warning label to any of the words in this category. An examination of the set of labels used by *OALD3* reveals a peculiarity. Compare the entries of *OALD3*:

> **harlot** *n* (archaic, or as a term of abuse) prostitute
> **tart**[3] *n* (derog sl) prostitute
> **whore** *n* ⚠ (derog) prostitute[6]

The list of labels found inside the front cover of *OALD3* does not name *term of abuse*. A possible explanation could be that *term of abuse* is not considered a label for "stylistic value" as the name for

---

[6] In examples such as the ones here, the aim is to copy the typographical conventions of each dictionary as closely to the original as possible. Following the policy of *OALD3*, the labels in these examples are typed in normal font, and not in italics, which is a common way of marking usage labels. However, in the text and in the tables of Appendix A, all labels are typed in italics for the sake of improved readability.

the list of labels reads, but is considered a part of the actual definition. However, the entries presented above show that the use of *term of abuse* is very similar to the labels specified in the user guide.

The calculations for frequencies of labelling in this word category can be seen in Table 1b. The dictionaries fall into three groups according to their relative frequencies of labelling. *OALD1*, naturally, stands apart from the rest by not attaching a negative label to any of the words it lists, thus receiving the relative frequency of 0%. *OALD2*, *OALD3* and *OALD6* form the second group by labelling 18%, 21% and 15% of the words, respectively. In *OALD4* and *OALD5*, labels are somewhat more common; they label 38% and 42% of the words listed.

Table 1b also shows the numbers of words marked as '+' and '(+)' separately. Of the five editions of *OALD* that assign cautionary labels to words for a prostitute, *OALD3* and *OALD6* show a slight preference for the simple type of label, whereas *OALD4* and *OALD5* represent the opposite. In *OALD2*, the numbers are equal. In this word group, a label is modified with a restrictive adverbial only once: the word *gigolo* carries the modified label *usu derog* in *OALD5*.

Alternative, inoffensive uses, on the other hand, are pointed out more often. A choice between a currency label and an attitudinal label is found, for instance, in the *OALD4* and *OALD5* entries for *whore*, which receive the labelling *dated or derog* in both dictionaries. Norri (2000: 88) criticizes this particular type of combination, as the two labels are not "mutually exclusive", and challenges the claim that the dated use is not derogatory (see Chapter 3.1 on page 16). The entries for *harlot* in *OALD2*, *OALD3* and *OALD4* provide further examples of similar combinations (*arch or term of abuse* in *OALD2* and *OALD3, arch or derog* in *OALD4*), thereby suggesting that such a use of the connector *or* has been a common practice in the Oxford line of learner dictionaries. Interestingly enough, in the entry for *harlot*, *OALD5* acts in line with Norri's view and labels the word as *arch derog*, omitting thus the connector *or*.

The entries for *scarlet woman* and *strumpet* in *OALD4* and *OALD5* also specify alternative situations of use. The word *scarlet woman* is labelled *dated derog or joc* in both dictionaries. Such a

string of labels may strike the reader as rather puzzling: is the label to be interpreted as [(dated derogative) or (jocular)] or [(dated derogative) or (dated jocular)]? In other words, is the jocular use of the word dated, too, or does the word in its jocular use belong to the contemporary vocabulary? Quite similarly, the label of *strumpet* (*arch or joc derog*) could be read either as [(archaic derogative) or (jocular derogative)] or as [(archaic neutral in attitude) or (jocular derogative, neutral in time)].

The "Detailed Guide" of *OALD4* (p. 1573) explains that usage labels may be combined, in which case they usually appear in the order: currency, region, register, evaluation (i.e. attitude), technical field and sayings. According to the classification of labels in *OALD4*, the labelling of *scarlet woman* is of structure 'currency evaluation or evaluation'. The labelling of *strumpet*, again, is of structure 'currency or evaluation evaluation'. The labelling clearly follows the policy set in the guide, but the explanation of the policy is not very helpful in interpreting such strings of labels.

A multiple cautionary label is used only once in this category, namely in the entry for *whore* in *OALD3*, which is labelled ⚠ *derog*. In *OALD3*, *whore* is simply defined as 'prostitute', but in the newer editions, the word receives two definitions: 'prostitute' and 'woman who has sex with many men'. The two separate definitions suggest that a process of semantic widening has taken place here. In *OALD4* and *OALD5*, the two definitions share the label *dated or derog*, while *OALD6* assigns the first sense the label *old-fashioned* and the latter sense a multiple description ⚠ "an offensive word….", thus implying that the original sense of the word has gone through a semantic amelioration.

When examining the likelihood of a particular word being treated similarly in all dictionary editions listing the word, sixteen of the twenty words are taken into consideration. Four words are left out, because they appear in one dictionary only. In such cases, a comparison to other editions is obviously not possible. In ten cases, a cautionary label is either attached or omitted consistently in all dictionaries listing the word. In nine of the cases, for *call-girl*, *courtesan*, *hooker*, *hustler*, *moll*, *prostitute*, *streetgirl*, *streetwalker* and *white slave*, the entries are unlabelled without

an exception. Conversely, *scrubber* is labelled in both dictionaries listing the word. Near-uniformity occurs in the labelling of *gigolo* and *harlot*, as one dictionary does not assign a warning label, while the others do this. As to the labelling of *tart*, the opposite is true – one dictionary labels it negatively, while other dictionaries listing the word do not. Labelling of *scarlet woman*, *strumpet* and *whore* is seen as non-uniform, meaning that there are at least two dictionaries to assign a negative label to the word, but also another two that leave out the label.

Taking a closer look at the use of the taboo sign, the dictionaries are quite harmonious in their views that there is no need for the taboo sign in this word category. Of the six words which receive a warning label in more than one dictionary (*gigolo*, *harlot*, *scarlet woman*, *scrubber*, *strumpet* and *whore*), *whore* is the only one to be labelled with the taboo sign, and this happens only in one of the six dictionaries (*OALD3*).

8.2 Treatment of Words for a Homosexual Person

Fourteen words referring to a homosexual person were included in the study. While most of the words in this category refer to homosexual men (such as *fag*/*faggot*, *fairy*, *pansy*), there are also words that are used of women only (*dike*/*dyke*, *lesbian*), as well as words that can be used of both sexes (*gay* and *homosexual*). The definitions of *gay* and *homosexual* have narrowed, however, toward a male referent only in the recent editions. Whereas the two words are defined as 'a homosexual person' in the earlier editions, the current definitions add the specification 'especially a man'.

The topic of this semantic field is, obviously, connected with sex and the taboo of propriety. In addition, homosexuality was once considered an illness and was illegal in many Western countries at the time when the earliest editions of *OALD* were current. Nowadays, the official and private attitudes to homosexuality have become more tolerant, but a certain feeling of unease still surrounds the topic: how to talk about homosexuality without being insensitive to the face wants of homosexual persons? Therefore, the topic involves the taboo of delicacy, too.

8.2.1 Inclusion

The number of words for a homosexual person found in the six dictionaries grows from zero (*OALD1*) to twelve (*OALD6*), as marked in Table 2b. As the total of the items investigated in this category is fourteen, the relative increase is considerable. The largest additions were made to *OALD2* and *OALD3*; while *OALD1* lists none of the fourteen items in this group, *OALD2* lists five of them. *OALD3*, again, doubles the number, resulting in ten of the fourteen terms. Since then, the increase has been more moderate, as the numerical data in Table 2b indicates.

As to uniformity in word inclusion, in the majority of cases (eleven of the fourteen words investigated) the word remains listed once it has been accepted into the dictionary. In only two cases, the word has been dropped out of the dictionary. These words are *closet queen* and *nancy*, both of which had a brief existence in the series by appearing in one edition only. The euphemistic use of *confirmed bachelor* in the sense 'homosexual man' is given in *OALD6* only; it is placed in the residue group.

8.2.2 Labelling

OALD3 is the earliest edition to assign labels indicative of a negative connotation to words for a homosexual person (Table 2a). The words that are negatively labelled in *OALD3* are not listed in the previous editions or are assigned attributes that in this study are regarded as belonging to other dimensions of language variation (e.g. *colloquial* and *modern use*). The number of different negative attributes is quite low in this word category, only four attributes are found altogether. Nevertheless, not any two dictionaries share exactly the same set of labels: △, *derog* and *offensive term* (*OALD3*); △ and *derog* (*OALD4*); *derog* and *offensive* (*OALD5*); and △, *disapproving* and *offensive* (*OALD6*). Again, the qualification "an offensive word" within the actual definition replaces the label *offensive* frequently in *OALD6*. In fact, the description is built into the definition seven times as opposed to one occasion of using the label *offensive*.

An interesting detail among the sets of labels is that *OALD5* does not attach the taboo sign, ⚠, to any of the words in this category, while *OALD4* and *OALD6* do this. Unlike the other dictionaries, *OALD4* does not attach *offensive* to any of the words for a homosexual person even though it has the label in its inventory.

Similarly as in the number of words listed, there is a clear change between *OALD2* and *OALD3* in their frequencies of labelling, as it can be seen in Table 2b. While *OALD2* does not attach a label of a negative connotation to any of the words it lists, *OALD3* labels 70% of its words in this group. Since then, the percentage has remained at roughly the same level.

Looking at the frequency of the taboo sign in particular, *OALD3* and *OALD6* use the taboo sign quite frequently in this word category (Table 2c). In *OALD3*, the taboo sign appears in all seven entries that are labelled negatively. *OALD6* assigns the taboo sign to six of eight negatively labelled words. *OALD4* uses the taboo sign as well, but considerably less often, only once of the seven entries which have a cautionary label or note.

As opposed to terms for a prostitute, in this semantic field it may quite safely be noted that the simple, unmodified type of label is preferred by all four dictionaries in which cautionary labels are used. While *OALD3* and *OALD4* use only simple labels, *OALD5* modifies the label of *dike*, suggesting that there are situations in which the word can be used without giving offence: *usu offensive*. *OALD6* also attaches the same modifier to *dyke* (*dike*), but quite curiously, uses also the taboo sign in the entry: ⚠ *usually offensive*. Such a string of labels seems somewhat contradictory, as the taboo sign is usually understood to indicate unconditional offensiveness. However, the list of labels and their meanings as explained inside the front cover of the dictionary offers a justification to the choice of labels. The taboo sign, ⚠, is explained as "likely to be thought by many people to be obscene or shocking" (*OALD6*, front cover), whereas the label *offensive* points to expressions that are "very insulting, especially in connection with their race, religion, sex or disabilities." The taboo sign can thus be interpreted as a warning of an offence against the decorum, while the label *offensive* refers more directly to offending a particular person or group of persons.

The word *queer* (in *OALD6*) is also marked as '(+)' in this study, but it differs from the entry for *dike* (*dyke*) in that it specifies the circumstances in which the word may be used without the offensive tone:

> **queer** *noun* (⚠, *slang*) an offensive word for a HOMOSEXUAL, especially a man which is, however, also used by some homosexuals about themselves

This dictionary entry shows thus one case of reclaiming a word, which means that the referents of the offensive word have begun to use the word in a positive sense as a sign of in-group membership. The word *queer* in *OALD6* is the only instance of specifying such an alternative, inoffensive, context in this semantic field.

As to using cautionary labels in combinations, *OALD3* and *OALD6* deviate from the others by favouring multiple labels. In all seven entries that are labelled negatively in *OALD3*, the label consists of the taboo sign combined to one or more cautionary labels. In *OALD6*, the share of words with multiple labels is a little lower, as seven of eight words receive a multiple label. *OALD5* assigns a similar combination to two of eight words, while *OALD4* does not use multiple labels at all in this word group. As to using cautionary labels in combinations, the points of interest are, on the one hand, linking the taboo sign, ⚠, to another label or note, and on the other, linking *disapproving* or *derogatory* to *offensive*. I will discuss the latter aspect first, and then move on to linking the taboo sign to another label.

Concerning the words in this semantic field, there are four instances of linking *derogatory* or *disapproving* and *offensive*. *OALD3* labels *pansy* with ⚠ *derog*, *offensive term*; *OALD5* labels *poof* (*poofter*) with *derog or offensive* and *queer* with *derog offensive* (without the conjunction *or*); and *OALD6* labels *fairy* with *disapproving*, accompanied by a qualification "an offensive word".

The fact that these multiple labels are not applied consistently to the four words in all editions of *OALD*, but rather are temporary occurrences, lends support to the claim made by Norri (2000) that the distinction between the labels *derogatory/disapproving* and *offensive* is not clear

(see Chapter 3.1, page 15). Moreover, the practice of *OALD5* of linking the two labels both with and without the conjunction reflects a hesitation in the policy.

Combinations of the taboo sign ⚠ and another negative attribute are found in this word group, as well. *OALD3* assigns the double label ⚠ *derog* to six words (*dike*, *fag*, *fairy*, *poof*, *queen* and *queer*) and the triple label ⚠ *derog offensive* to the word *pansy*, as it was already noted above. *OALD6* combines the taboo sign ⚠ and the description "an offensive word" five times (*fag*, *pansy*, *poof*, *queen*, *queer*), and once ⚠ *usually offensive* (*offensive* as a label, in the entry for *dike*).

As to the likelihood of similar treatment in labelling in the dictionaries studied, eleven words of the total of fourteen in this semantic field could be evaluated. This is because three of the words appear only in one dictionary (*closet queen*, *confirmed bachelor* and *nancy*). In the eleven cases where the word was listed in more than one dictionary, the editors are quite united in their labelling. In nine cases, the dictionaries either consistently apply a cautionary label to the word or consistently do not apply one. Words furnished with a label in all dictionaries are *dike, fag, fairy, ponce, poof* and *queen*. Words without a warning label in all dictionaries are *gay*, *homosexual* and *lesbian*. It could then assumed that the set of words labelled in all dictionaries are the most likely to offend in this semantic field, while the latter ones constitute a neutral and stable core vocabulary. In the remaining two cases (*pansy* and *queer*), one dictionary is in dissonance with the others, as the words do not have a cautionary label in *OALD2*.

On the need for the taboo sign in particular, the dictionary editors agree to a lesser extent. Of the eight words which receive a cautionary label in more than one dictionary, only one word (*ponce*) is treated uniformly; the sense 'homosexual' is not furnished with the taboo sign in neither of the two dictionaries giving the sense. The word *dyke* is not assigned the taboo sign in *OALD5* whereas in the other dictionaries the sign is used; the converse holds for *fairy* in *OALD3*. These two words make up the category of near-uniformity. The category of non-uniformity forms the majority, as five of the eight words (*fag*, *pansy*, *poof*, *queen* and *queer*) are marked with the taboo sign in *OALD3* and *OALD6*, but not in *OALD4* and *OALD5*.

8.3 Treatment of Words for Having Sex

The forty-two words in this category are not all true synonyms to each other; rather, the category could be characterized as comprising of near-synonyms and partial synonyms which share the meaning components 'have a relationship that involves sexual acts' and 'human agent'. When making decisions about whether a particular word should be admitted to the analysis, I have relied on the definitions given in the dictionaries. If any one of the six dictionaries indicated that the word in question has the two meaning components, it was taken in.

The first of these meaning components, 'have a relationship that involves sexual acts', allows that the words range from ones that refer explicitly to the act of having sex (*ball*, *bang*, *bonk*, etc.) to ones that emphasize the relationship between the persons (*be intimate*, *commit adultery*, *consummate a marriage*, etc.).

The meaning component 'human agent' is intended to rule out words that are used of animals only. However, if any one of the dictionaries indicates that the word can be used of both humans and animals, the word is taken in. Consequently, the word *couple* is regarded as being not listed in *OALD1* to *OALD3*, but as listed in *OALD4* and the newer editions. Compare the entries of *couple* in *OALD3* and *OALD4*:

> **couple** *vt, vi* […] **2** marry; (of animals) unite sexually, (of things) come together; unite. (*OALD3*)
>
> **couple** *v* […] **3** (*arch or rhet*) have sexual intercourse. (*OALD4*)

As *OALD3* defines the verb *couple* as 'marry' when used of human beings, the sense 'have sex' is regarded as 'not listed'. Of course, it could be claimed the meaning component 'having a relationship that involves sex' belongs inherently to the verb *marry*, which is found in the definition of *couple* in *OALD3*. The example shows that it is not always easy to draw a distinct line between semantic fields, as the borderline areas are often fuzzy and overlapping. In this study, judgements about whether a word belongs to a particular semantic field or not are largely based on the wording of the dictionary definition.

A number of the items in this semantic field consist of a verb and a noun or an adjective (for instance, *commit adultery, have sex, have intercourse, have carnal knowledge* and *be intimate*). As the verb is "empty" in meaning, the entry for the noun or the adjective is investigated.

As to the taboo type, words for having sex are primarily taboos of propriety. It is a matter of knowing the correct words for each context, in order to keep a positive image of oneself in the eyes of other people.

8.3.1 Inclusion

Similarly to in the previous word category, the number of words for having sex grows quite remarkably from the oldest edition to the newest one. As indicated in Table 3b, *OALD1* lists only two of the forty-two words investigated, whereas the equivalent number is thirty-eight in *OALD6*. Every edition contains a higher number of terms than its predecessor; the most notable addition is made in *OALD4*, in which the number of words rises by fourteen. Conversely, the increase in the number of words is at its lowest in the next edition, in *OALD5*, when the number goes up by one word only.

As to the consistency of word inclusion (as seen from Table 3a), most words remain listed once they have made their way in. *Commit adultery* is the only verb listed in all dictionaries and thirty items (equalling 71%) have made their entrance later and have been listed ever since. Four items (*bang*, *knock up*, *make* and *take*) fluctuate between being listed and left out, while *lie with* is the only verb to disappear for good. *OALD6* adds six words (*ball*, *do it*, *get it on*, *get your leg over*, *go all the way* and *poke*), but deletes none; these additions belong to the residue group.

8.3.2 Labelling

The attributes used to mark words with negative connotations in this category include taboo sign △ and *vulg* (*OALD3*); △, *derog*, and *sexist* (*OALD4*); △ and *derog* (*OALD5*); △ and *disapproving*

(*OALD6*). As it can be seen from Table 3a, the oldest editions, *OALD1* and *OALD2*, do not label any of the verbs for having sex.

A point of interest in the range of labels is *sexist*, used by *OALD4*. Unfortunately, *OALD4* uses the label only once in this word group, for the transitive use of the verb *make* "succeed in having sex with (a woman)", which is not included in any of the other dictionaries. Therefore, the use of *sexist* cannot be compared to labels of the other dictionaries. However, a closer look at a few entries in *OALD4* shows that the use of the label *sexist* is not completely straightforward to the dictionary makers either:

> **make** *v* **18** (*sl sexist*) succeed in having sex (with a woman): *The guy doesn't make the girl until the last chapter.*
>
> **have³** *v* **10** (⚠ *sl*) (esp of a man) have sexual intercourse with (sb): *Have you had her yet?*
>
> **take** *v* **36** (of a man) have sexual intercourse with (a woman): *He took her on the sofa.*

Even though the meanings and contexts of use for *make*, *have* and *take* are not exactly the same, the example sentences in all three entries suggest, in my view, a similar attitude towards the person (i.e. usually a woman) talked about; she is seen as the object of actions taken by a man. Assuming that this is the reason behind the label *sexist* for *make*, one wonders why the other two entries lack the label.

OALD1 and *OALD2* label none of the words they list in this semantic field, as already pointed out. *OALD3* assigns a cautionary label to 29% of words listed and *OALD4* reaches the peak by 35% (Table 3b). Conversely, *OALD5* uses cautionary labels on 22% of the items, which is the lowest figure in the newer editions. The taboo sign ⚠ is used often in this semantic group. *OALD3* attaches the sign to all of its negatively labelled words, and the following editions to the majority (Table 3c).

The dictionaries clearly prefer unmodified labels for verbs for having sex. The word *fornicate* is the only word to receive a modified label, namely *fml esp derog* in *OALD3* and *fml usu derog* in *OALD4*. The modifier *esp* may appear quite ambiguous to language learners; a potential,

but not the intended, interpretation could be 'formal, very derogative'. For this reason, the change to *usu derog* in *OALD4* is clearly for the better.

Labels indicative of a negative attitude are combined into a multiple label only once in this category; *OALD3* describes the verb *knock up* "(of a man) have sexual intercourse with; make pregnant" as both △ and *vulg*. By comparison, *OALD4* and *OALD5* attach only the taboo sign to the entry for *knock up*, but these entries are marked as '0' in this study, because the definition carries only the narrower meaning 'make pregnant'.

Another combination of labels to draw attention to is the verb *roger* (in *OALD4*), which is assigned both the labels △ and *euph*. Of course, the latter label is regarded to be outside the scope of this study, because it is seen more as a positive label, meaning 'use this instead of *x*'. If the meaning of the label is such, and if the meaning of the taboo sign is 'avoid this as it may offend', then a combination of the two does indeed represent a contradiction. It must be kept in mind, however, that the two labels are closely connected, as euphemisms can serve as substitutes for tabooed words. From a sociolinguistic point of view, a more exact labelling would perhaps be △ *dysphemism*, but dictionaries quite understandably want to narrow down the number of different labels. In *OALD5* and *OALD6*, the cautionary label of *roger* is simply △.

As to the consistency of labelling, eight of the forty-two items are not comparable, as they are listed in one dictionary only. Twenty-three items are unlabelled in all works listing the item and another five items (*fuck*, *hump*, *roger*, *screw* and *shag*) are labelled in all works that list the word. In other words, twenty-eight verbs (82%) are labelled in a uniform manner. The verbs *fornicate*, *have* and *lay/get laid* belong to the category of near-uniformity, while *bang*, *have it away* and *have it off* are treated more diversely.

Nine verbs are assigned a negative label in more than one dictionary; they are taken into consideration when calculating the consistency of assigning the taboo sign. While the verb *fornicate* never receives the taboo sign, the remaining eight verbs (*fuck*, *have it away*, *have it off*, *hump*, *lay* as

in *get laid*, *roger*, *screw*, *shag*) are assigned the taboo sign throughout. In other words, the consistency of assigning the taboo sign to items with a cautionary marking is 100%.

8.4 Treatment of Words for Buttocks

The borders of this semantic field appear not as fuzzy as in some other fields of this study. Many of the words in this group are defined simply as 'buttocks' or 'the part of the body on which one sits'. To put it differently, it was quite easy to determine which words to take into investigation and which to leave out.

Being words that refer to a private body part, this semantic field belong to taboos of propriety. Obviously, the degree of controversy is quite mild in comparison to, for instance, words for having sex. Thus, it is interesting to see how the treatment of bodily taboos differ from the treatment of sexual taboos.

8.4.1 Inclusion

The numbers of words for buttocks found in the dictionaries (Table 4b) present quite a significant increase from five words of *OALD1* to seventeen of *OALD6* of the nineteen words. Similarly as in the semantic field of 'prostitute', *OALD5* deviates from the other editions by listing fewer words than its predecessor. Turning the figures into percentages, the inclusion rates are 26% in *OALD1*, 47% in *OALD2*, 58% in *OALD3*, 84% in *OALD4*, 79% in *OALD5* and 89% in *OALD6*. Thus, the largest additions of words are found in the earliest editions.

Fifteen of the nineteen items (79%) in this semantic field are kept listed once they have been introduced for the first time. Five of these (*behind*, *bottom*, *buttocks*, *rump*, *seat*) are listed in every edition, while ten items emerge later. None of the words seems to have disappeared from the vocabulary, but two words fluctuate (*stern* and *tail*). Another two items were added by *OALD6* (*buns*, *duff*); their treatment on the long run remains to be seen.

8.4.2 Labelling

Only three of nineteen items receive cautionary markings in this semantic field (Table 4a). These are *arse*, *ass* and *backside*. *OALD1* does not assign labels to any of its five words in this semantic field. *OALD2* describes *arse* as ⚠, *not in polite use* and *backside* as *not used in polite society*. The word *ass* is not listed in *OALD2*. *OALD3* attaches the multiple description ⚠, *not in polite use* to *arse* and the label *vulg* to *ass*. From *OALD4* on, both *arse* and *ass* are furnished with the taboo sign, ⚠, in all three editions. To put it short, the frequencies of labelling vary between 12%-22%: every dictionary except *OALD1* assigns a cautionary label on two words, but as the numbers of words included differ, the percentages differ also. The dictionaries prefer unmodified labels. Multiple labels as combinations of the taboo sign and another attribute are found in *OALD2* and *OALD3*, whereas the newer editions use the taboo sign as a simple label. As to the uniformity of labelling, *arse* and *ass* are labelled harmoniously, but on the labelling of *backside*, one dictionary deviates from the rest.

There are two points worth noting. First, *OALD3* does not use the taboo sign on *ass*, even though it uses the sign on *arse* and also quite frequently in other semantic fields. Second, *OALD2* has two labels that have quite similar meanings, and that are close in their form, as well: *not in polite use* and *not used in polite society*. This is perhaps an indication of dictionary making not yet being a very systematized and fixed business at that time.

As the labels are few in this semantic field, it is perhaps interesting to look at other attributes assigned to words in this group (Table 4a). The only item without any label whatsoever in the line of dictionaries is *buttocks*. Quite a few items receive the attributes *jocular* or *humorous*, some are also pointed out to be *euphemisms*. Labels for formality and style are also used: for instance, *formal*, *informal*, *colloquial* and *slang*. Further, regional restrictions are occasionally indicated: *BrE* and *AmE*.

8.5 Treatment of Exclamations of Anger

Words used as exclamations originate from a variety of word classes; most often they are derived from nouns (for example, *ballocks*, *shit*, *hell*, *Christ*) and verbs (*fuck*, *damn*, *sod*). There are also items that do not have a distinct denotative meaning, but are coined solely for the purpose of expressing a feeling (*darn, dash, heck, shite*). However, these items are not entirely without a denotative content, either. They are often euphemistic alterations of nouns and verbs used as exclamations, as the entries indicate:

> **heck** *n* (sl, euphem) hell (used in exclamations): *Oh! What the ~!* (*OALD3*)
>
> **shite** *exclamation, noun* [U] (*BrE*, △, *slang*) another word for SHIT (*OALD6*)

Meanings of verbs and nouns used as exclamations often involve topics of sex (such as *fuck*, *bugger*), body and its functions (*shit*, *bollocks*) and religion (*damn*, *God*, *Christ*, *hell*). Ljung (1984: 29-30) observes that religious swearwords have lost much of their power in at least the Swedish language, as religious commitment has diminished, but implies that this may very well be the case in other Western languages, too. However, the writer admits (1984: 30) that some religious oaths are still regarded vulgar, and the use of such religious as well as sexual and bodily swearwords is considered a violation of good taste and manners; the shock value is not in the literal meaning, but in the speaker's will to use the vulgar words.

Following Ljung's line of thought, the taboo types connected with this word group used to be both taboo of fear and taboo of propriety. Nowadays, the latter type is probably dominant.

8.5.1 Inclusion

The policy of deciding which words qualify as 'listed' in this word group is somewhat looser than in the other groups. This was found necessary, because words used as exclamations are not always marked as exclamations or interjections in the dictionaries, but the expressive function is sometimes explained in entries of items marked as verbs or nouns. If the example sentence or any other part of the entry implies that the item can be used as an exclamation, the word is considered as listed.

Altogether thirty-five exclamations of anger were taken into investigation. The number of exclamations found in the dictionaries increases quite significantly from nine items of *OALD1* to thirty-four of *OALD6*, as Table 5b tells. That *OALD6* receives such a high figure is due to the fact that the dictionary served as the main source of the original word list, which was already predicted in Chapter 5.2 (p. 31). In a diachronic perspective, the increase in the number of exclamations is quite steady, except in *OALD5*, which deviates from the general trend by keeping the number the same as in the previous edition. All the other editions made additions to their coverage of exclamations of anger.

Having *OALD6* as the source of the word list has its consequences on the numbers of consistency, too; there are obviously no words that seem to have disappeared from the vocabulary. Thus the categories of consistency that give any reliable data are the eight items that appear in all editions of *OALD* (*bother*, *for Christ's sake*, *confound (it)*, *damn (it)*, *dear*, *Heavens*, *hell* and *Lord*) and the two items which fluctuate between being listed and left out (*flip* and *God*). The first set of exclamations form a stable, core vocabulary of the field. The reason for the fluctuating status of the latter two items remains unexplained.

8.5.2 Labelling

The descriptions used to indicate the hazards of using exclamations are as follows: $\triangle$, *vulgar*, *not in polite use* and *term of abuse* (*OALD2*); $\triangle$, *vulg* and *not in polite use* (*OALD3*); $\triangle$ (*OALD4* and *OALD5*); and $\triangle$ and *offensive* (*OALD6*). As already pointed out, *OALD6* often incorporates a usage note in the definition. The attribute is *offensive* in all of the cases, but the form of the description varies somewhat, which can be seen from the following examples of *OALD6*:

hell *noun* **3** a swear word that some people use when they are annoyed or surprised or to emphasize sth. Its use is offensive to some people: *Oh hell, I've burned the pan.* ° *What the hell do you think you are doing?* ° *Go to hell!* ° *I can't really afford it, but, what the hell* (= it doesn't matter), *I'll get it anyway.* ° *He's as guilty as hell.* ° *(AmE)* '*Do you understand?*' '*Hell, no. I don't.*'

fuck *verb* (⚠, *slang*) **2** a swear word that many people find offensive that is used to express anger, disgust or surprise: *Oh, fuck! I've lost my keys.* ° *Fuck it! We've missed the train.* ° *Fuck you—I'm leaving.*

ball *noun* (⚠, *slang*) **3** (**Balls!**) *(BrE)* *exclamation* used as a swear word when you are disagreeing with sth, or when you are angry about sth **HELP** Less offensive ways to express this are '**Nonsense!**', or '**Come off it!**'

The example entries show that the built-in usage indication sometimes stands alone without a separate label (as in the entry for *hell*), while in other cases, it enhances the force of the taboo sign (as in the entry for *fuck*) or gives advice how to avoid the taboo (as in the entry for *ball*). The dictionary also appears to make a distinction between the descriptions "offensive to some people" (*hell*) and "offensive to many people" (*fuck*). Interestingly enough, the entries that contain the description "offensive to many people" are also furnished with the taboo sign, whereas the description "offensive to some people" is never combined to the taboo sign in this word group. In other words, "many people" correlates with the taboo sign[7].

As the calculations of Table 5b show, the frequencies of cautionary labelling vary between 0% and 35%. *OALD1* lists nine exclamations, but labels none of them. *OALD2* assigns a label to 25% of its exclamations, from which the frequencies rise to 35% of *OALD6*. The figures show that the rise is not steady, but keeping in mind the small number of entries investigated, the slight fall of the frequencies in *OALD4* and *OALD5* is not important.

The use of the taboo sign ⚠ is frequent in this word group Table 5c). *OALD4* and *OALD5* attach the sign to all negatively labelled words, and the other dictionaries do the same on most words. *OALD2* and *OALD6* prefer to use the taboo sign combined to another cautionary label, whereas *OALD3* shows a slight preference for the opposite. *OALD4* and *OALD5* represent an

---

[7] Dictionary entries which contain the taboo sign combined to the description "offensive to many people" are labelled as '+', even though the description implies that there are people who do not find the exclamation offensive. The taboo sign is seen as overriding the verbal description in such cases.

extreme preference for the taboo sign, as they do not use any other cautionary descriptions in this word group.

All multiple cautionary markings in this category contain the taboo sign ⚠. In *OALD2*, *shit* (⚠, *not in polite use*), *bugger* and *sod* (both ⚠, *vulgar term of abuse*) are assigned multiple descriptions. *OALD3* uses a multiple label twice, for *ballocks* (⚠, *vulg*) and *shit* (⚠, *vulg*, *not in polite use*). *OALD6* combines ⚠ and *offensive* six times (*balls*, *Christ*, *fuck*, *shit*, *sod*, *son of a bitch*); *offensive* is found in the definition part all the times, but the expression varies: for instance, a help note in the entry for *balls* indicates offensiveness implicitly: "Less offensive ways to express this are '**Nonsense!**', or '**Come off it!**'" The entry for *Christ* reads "a swearword that many people find offensive…." Yet another way of indicating the same can be found in the entry for *son of a bitch*: "an offensive word…."

Overall, unmodified labels are preferred to modified ones in entries for exclamations. A modifier is used once, in the entry of *shit* in *OALD5*, which is marked *usu* ⚠. This supports the view that the taboo sign does not indicate an unconditional prohibition of use, but can be overrun like any other type of label. The three entries of *OALD6* which hold the description "offensive to some people" are also marked as modified cautionary markings, because the description implies is that not all people find the given item offensive.

The dictionaries are quite harmonious in their labelling. Of the thirty exclamations that can be compared in this respect, twenty-six are labelled similarly in all dictionaries, equalling 87%. On three items (*for Christ's sake*, *God* and *hell*), one dictionary acts differently by attaching a negative label while the others do not. Greater diversity is found only in the entries of *blimey*.

On the use of the taboo sign in particular, the consistency is even stronger. Eight exclamations that are labelled negatively in more than one dictionary receive the taboo sign in all of those dictionaries (*ballocks/bollocks*, *balls*, *bugger*, *Christ*, *fuck*, *shit*, *sod* and *son of the bitch*). Meanwhile, one item (*blimey*) does not receive the taboo sign in neither of the dictionaries that label it negatively.

It is rather surprising that this word group contains quite a number of items that are not assigned a cautionary label. Many of the unlabelled exclamations are religious words (such as *damn*, *Heavens* and *Lord*) or their alterations (*crikey*, *darn*, *heck*) or alterations of bodily or sexual swearwords (*flip*, *shoot*, *sugar*). These are often marked as euphemisms and it remains to be seen whether these become tainted by the original term and will be labelled negatively in the future. By comparison, the exclamation *shite* (appearing in *OALD6* only) is assigned the taboo sign.

8.6 Treatment of Words for a Person Belonging to an Ethnic Minority

This word group consists mainly of names for persons belonging to two ethnic minority groups, namely black people – or African Americans, which seem to be the current neutral term in American English – and Native Americans. A few terms with a wider range of application are included as well. Consequently, the items in this group do not have a common denotation, but they all share the meaning feature 'a member of an ethnic minority'. While some words (for instance, *wog*, *non-white* and *native*) have a more general meaning, other ones (*boy, brave, buck, mammy, piccaninny* and *squaw*, among others) represent the opposite with a narrower range of possible referents.

As politically correct language is taken for granted in all public communication nowadays, paying attention to terms for representatives of ethnic groups is worthwhile. This area of vocabulary is linked to taboos of delicacy as well as to taboos of propriety. A foreign learner may, for example, wonder how to refer to persons of different ethnic backgrounds without sounding offensive and unwittingly hurting the feelings of those persons. Another concern is more egocentric: how to refer to these people without sounding ill-mannered and ignorant of the current social standards of political correctness?

8.6.1 Inclusion

Table 6b presents a numerical summary of the treatment of names for a member of ethnic minority. The number of words included in each dictionary ranges between nineteen (*OALD1*) and twenty-eight (*OALD4*) of the thirty-four. Viewing chronologically, every new edition brings a slight increase in the number of words listed until *OALD5*, which goes down by four words in comparison to the previous edition. *OALD6* increases the number again, but it does not reach the peak of *OALD4*.

When the inclusion of individual words is considered diachronically, one word (*boy*) must be ignored, as it does not appear in the relevant meaning until in *OALD6*. In other words, its fate remains yet to be seen. Fourteen of the thirty-four words are listed in all six editions, another ten remain listed once they are taken in. In other words, 71% of the words in this category remain listed once they have been listed for the first time. There are, nevertheless, cases of the opposite treatment. Four words are judged as having disappeared from the dictionaries (*blackamoor*, *Jim Crow*, *mammy* and *piccaninny/pickaninny*), as they have not been listed in the relevant sense in at least the last two dictionaries. Another five words are marked as fluctuating between being listed and left out again (*Afro-American*, *buck*, *darkey/darkie/darky*, *Negroid/negroid*, *person/man/woman of colour*).

8.6.2 Labelling

The attributes describing negative connotations in names for an ethnic minority person are the following: *contemptuous* and *derogatory* (*OALD1*); *contemptuous*, *derog.*, *impolite* and ⚠ (*OALD2*); *contemptuous*, *derog*, *offensive*, *impolite* and ⚠ (*OALD3*); *derog*, *offensive* and ⚠ (*OALD4* and *OALD5*); *offensive* and ⚠ (*OALD6*). *OALD3* uses the widest set of attributes, but quite oddly, three of the five attributes used as labels (*contemptuous*, *offensive* and *impolite*) are not included in the list of style labels printed inside the front cover of the work.

Calculations for the relative frequencies of labels or equivalent comments can be found in Table 6b. The range of frequencies is quite wide, continually increasing from 11% of *OALD1* to

73% of *OALD6*. However, the range appearing wide does not mean that labels in the first edition are rare in this word group in comparison to the other word groups. Rather on the contrary, this is the only word group in which *OALD1* uses cautionary labels. The dictionaries can be divided into three groups as to their eagerness to use labels. *OALD1* and *OALD2* use labels on a little over 10% of the words they list, *OALD3* on roughly 40%, and finally, the newest editions (*OALD4* and on) on more than 60% of their entries.

Examining the frequency of using the special taboo sign in particular, the following points can be pointed out (Table 6c). Except for the oldest one, *OALD1*, the dictionaries have the taboo sign in their inventories of cautionary labels. In this word category, the sign is used by all these five dictionaries. The taboo sign is never used as a single label in this category, but appears always combined to another label. *OALD2* attaches the taboo sign to one word of the three words (33%) which it labels negatively. For the other editions, the frequencies are 60% (*OALD3*); 35% (*OALD4*); 33% (*OALD5*); 47% (*OALD6*). The absolute numbers are quite low, but they are turned into percentages for the sake of comparison. *OALD3* is the most eager of the dictionaries to attach the taboo sign to words of this category, and *OALD6* comes closest behind. *OALD2*, *OALD4* and *OALD5* are all roughly at the same level as to the frequency of the taboo sign.

Table 6b also shows separately the numbers of words marked with an unmodified label words with a modified label, or a label specifying an alternative context of use. Both types of labels, modified as well as unmodified ones, are used in every dictionary. *OALD4* and *OALD6* show a more visible preference for unmodified labels, while for the rest the numbers are equal or the difference is of one instance only. *OALD5* differs from the other dictionaries by expressing a slight preference for modified labels; seven cases of unmodified labels against eight cases of modified ones. Among the adverbials used as modifiers are *esp* (for *especially*), *sometimes*, *often* and *usu* (for *usually*). The last three modifiers clearly restrict the range of application of the label, suggesting that there are situations in which the word can be used without the negative connotation.

Norri (2000: 79) observes that the choice of the modifying adverbial seems to "fairly arbitrary". One might also criticize the modifiers from the point of view of a learner; the difference is not always clear. For instance, *OALD5* uses all three modifiers *sometimes, often, usu* (for *usually*) – making a distinction between them may not be an easy task for a language learner.

The meaning of *esp* appears somewhat ambiguous, as already discussed in Chapter 8.3.2. (on page 58). A dictionary user may be puzzled by the label *esp offensive* (for *Native/native*), for instance. Do the dictionary editors wish to indicate that the expression is 'very offensive' or 'used especially in situations where the speaker wishes to be offensive'? An examination of other combinations in which the modifier *esp* occurs shows that the ambiguity is not always present. The most usual type of label to which *esp* is attached appears to be regional labels. In such cases the interpretation is unproblematic; in the entry found in *OALD4* for *brolly* 'umbrella' (labelled *infml esp Brit*), the meaning of the label is undoubtedly 'used especially in British English', not 'the word is very British'. The same applies for the entry for *brownie* 'cake', which is assigned the label *esp US* in *OALD4*. The possibility of ambiguity quite obviously depends on the gradability of the adjective used as a label; combined to gradable adjectives (such as *offensive*), both interpretations of *esp* are at least theoretically possible. If combined to non-gradable adjectives, there seems to be only one possible meaning for *esp*.

OALD6 uses yet another type of modifier in its cautionary notes, *very*, but these entries are not marked as '(+)', because the adverbial does not limit the range of application of the label or comment. Conversely, it strengthens the force of the label. The words to which the comment *very offensive* is attached are *coon, nigger, Red Indian* and *wog* in *OALD6*.

The entries also contain modified labels that suggest a change in the status of the word. In this word group, such instances of labelling are found in *OALD3*, *OALD4* and *OALD6*. For example, the word *mammy* receives the label *dated now offensive* for in *OALD4*. I have treated these temporal modifiers similarly to the ones above (*esp, sometimes, often, usu*), regarding them as limiting the range of application of the label. This procedure seems warranted, as the label implies

that the learner may still meet the word in contexts where the label does not apply. From a strictly synchronic point of view, the meaning of such a label is, of course, unmodified.

Following the line of dictionaries, such a change also seems to have taken place in the status of the noun *black/Black*. The word appears unlabelled in *OALD1* and *OALD2*. In the subsequent editions, the strings of labels are as follows: *formerly derog, but now widely used* (*OALD3*); *formerly derog, now the preferred word* (*OALD4*); *sometimes offensive* (*OALD5*). In *OALD6*, the word is unlabelled, but the entry contains a note: "In this meaning **black** is more common in the plural. It can sound offensive in the singular. Instead, you can use the adjective ('a black man/woman') or, in the US, **African American**." Thus, the oldest two dictionaries indicate a neutral status, whereas the next two editions claim that the word has had a negative tone which is now lost. The latest two editions indicate a conditional negative connotation, for a change. There has apparently been a change in the status of the word, but the dictionaries are contradictory as to when the change has taken place.

Alternative contexts of use are specified four times in this word group. The word *blackamoor* is labelled *hum or derog* in *OALD2* and *OALD3*. *OALD6* labels both the nominal and adjectival uses of *coloured* (i.e. the terms *coloured* and *coloured person*) as *old-fashioned or offensive*. The latter cases present, again, an example of alternatives that are criticized for not excluding each other. However, some justification can be expressed on this particular case. The chronological line of labels for *coloured person* shows that the term remains unlabelled from *OALD1* to *OALD3*. *OALD4* uses a currency label *becoming dated*, while *OALD5* describes the term as *dated often offensive*. Thus, there seems to have been a process of semantic pejoration or tainting going on. If that is the case, both attributes, *offensive* and *old-fashioned*, do describe the term accurately, but something could perhaps be made to render the labelling more exact. Instead of pairing the two labels as alternatives to each other, combining them as *old-fashioned, now offensive* would possibly describe better the current status of the word.

Multiple descriptions are numerous in this word category. In fact, *OALD1* is the only dictionary not to use multiple labels. For the other works, the number of multiple labels against all instances of labelling are as follows: one multiple label against three instances of labelling in total (*OALD2*); 6/10 (*OALD3*); 8/17 (*OALD4*); 5/15 (*OALD5*); 9/19 (*OALD6*). All but two instances of multiple labelling involve a combination of the taboo sign and another attribute: ⚠ *impolite*, ⚠ *derog*, ⚠ *offensive*, ⚠ *contemptuous*, ⚠ *now derog*, ⚠ *very offensive*, ⚠ *impolite offensive* and ⚠ *derog offensive*. The combination *derog offensive* (without the taboo sign) is found twice in this word group (*OALD4* for *blackamoor* and *Jim Crow*).

The labelling of the word *nigger* provides an interesting example of variation in cautionary labelling: ⚠ *impolite word* (*OALD2*); ⚠ *impolite and offensive word* (*OALD3*); ⚠ *derog offensive* (*OALD4* and *OALD5*); ⚠ "a very offensive word…" (*OALD6*). The word *nigger* appears also in *OALD1*, but without a label for a negative connotation: it is labelled *colloq*. The taboo sign appears to be a more permanent label than the other attributes assigned to the word. The other labels are perhaps another piece of evidence of the fact that dictionaries do not always have clear distinctions between different cautionary labels.

Looking at the consistency of labelling of individual terms across the dictionaries, two terms (*boy* and *buck*) are ignored because they appear in one dictionary only. As a result, thirty-two terms are investigated in this respect. Thirteen of the terms (41%) are treated uniformly in all dictionaries listing the word. In nine cases (*African American, Afro-American, American Indian, Amerindian, brave, Native American, Negroid, non-white* and *person of colour*) the term is unlabelled, whereas four terms (*coloured, Jim Crow, Uncle Tom* and *wog*) are labelled in all dictionaries covering the term. On the labelling of eight terms (25%), one dictionary is in disharmony with the rest. In total, the groups of uniformity and near-uniformity make up 66% of the words in this category, while the rest are treated in a more varied manner.

Of the thirty-four words in this category, nineteen receive a negative label in more than one dictionary. These words are taken into account when analysing the uniformity in the use of the

taboo sign between the dictionaries. On eight of the nineteen words, none of the dictionaries use the taboo sign as the cautionary marking (*black, black person, blackamoor, coloured, coloured person, native, Negress* and *Negro*). Conversely, on four words (*coon, darkey, nigger,* and *wog*) the dictionaries unanimously use the taboo sign as a cautionary marking. On five words, one dictionary stands apart from the rest in their judgement on the need of a taboo sign. *OALD6* adds the taboo sign to the negative labelling on *half-breed* and *half-caste*, while the others leave it out. *OALD3* does the same on *Jim Crow*. Oddly enough, while *OALD3* appears quite eager to use the taboo sign in comparison to the other dictionaries, it stands out by not assigning one to *Red Indian* and *redskin*. A potential reason can, however, be found by following the chronological line of labelling of these words. In *OALD1* and *OALD2*, *Red Indian* and *redskin* appear unlabelled; in *OALD3,* both receive the labelling *old use, now impolite*. In the newer editions, the words are assigned a multiple label or an equivalent comment, the attributes being the taboo sign, $\triangle$, and *offensive*. These labels may in fact reflect a change in the language; perhaps the editors of *OALD3* did not use the taboo sign because the degree of offensiveness was not yet considered very high at that time.

8.7 Treatment of Words for a Stupid Person

This semantic field contains many figurative and colourful expressions. Some of the words suggest that the brain consists of inactive material (*blockhead*, *bonehead*, *thickhead*) or is empty (such as *airhead*, *dunderhead*). Another set of words equates a human being with an animal; for example, *ass*, *jackass*, *birdbrain*, *donkey*, *muttonhead* and *mutt*. *Cretin*, *idiot*, *imbecile* and *moron* are examples of semantic widening in which attributes formerly used in medical contexts have been taken into everyday language with a more general meaning. Three words in this group – *dickhead*, *twat* and *wanker* – have a connection with the taboos of body and sex. The first of these, *dickhead*, is perhaps an alteration of the words ending in '*head*' listed above. Similarly, the *OED* indicates that *twat* is an alteration of *twit* (*OED Online*, s.v. *twat*). Thus, *twat*, *wanker* and *dickhead* share the

fact that they are not only taboos of delicacy, but belong to taboos of propriety by their literal meaning.

Quite understandably, it is not polite to say that a person is not intelligent. Consequently, items in this semantic category are controversial for reasons of delicacy, primarily. However, it is also a question of propriety, because these words are often used with the intention of insulting. The taboo of propriety is naturally a dominant factor for the three words with the literal meaning connected with body and sex.

8.7.1 Inclusion

Table 7b shows the number of words for a stupid person found in the six dictionaries. The numbers range between twenty-four (*OALD1*) and forty-nine (*OALD4* and *OALD6*) of the sixty-two items in this semantic field. Turning the figures into percentages, the range of words included is from 39% to 79%. Following the line of dictionaries chronologically, the number of words shows quite a steady increase until reaching the peak in *OALD4*. As in many other word groups, *OALD5* brings a temporary decline in the number of words, which is brought back up in the sixth edition.

As to the consistency in which individual items are kept listed, fourteen words appear in all six editions, another twenty-four remain listed after they have been covered for the first time. Put together, these for 62% of the group. Five words seem to have been left out. Quite a considerable number, eleven words (18%), fluctuate between being listed and left out. The residue group consists of six new words and two words that are dropped out by *OALD6*.

8.7.2 Labelling

Words in this group are first labelled in *OALD4*. As in many of the previous categories, *OALD4* and *OALD5* use the same set of labels: ⚠, *derog* and *offensive*. *OALD6* has the set of labels ⚠, *disapproving* and *offensive*. In qualifications within the definition, *OALD6* introduces a new adjective, *rude / very rude*.

As Table 7b shows, *OALD4* assigns a cautionary label to fifteen of forty-nine words (equalling 31%), *OALD5* to thirteen of forty-two (31%) and *OALD6* to ten of forty-nine (20%). The taboo sign is used in *OALD4* in three of fifteen negatively marked entries (*cretin*, *twat* and *wanker*), *OALD5* in two of thirteen (*twat* and *wanker*) and *OALD6* in three of ten entries (*dickhead*, *twat* and *wanker*). Following each of these entries chronologically, *cretin* is assigned a multiple label ⚠, *offensive* in OALD4, but *offensive* only in OALD5 and OALD6. *Dickhead* appears in the newest two editions only, receiving the marks *derog* (in *OALD5*) and ⚠ and "a very rude way of referring…" (in *OALD6*). *Twat* and *wanker* receive the same sets of attributes: ⚠, *derog* for both words in *OALD4* and *OALD5,* and ⚠ combined to the attribute *offensive* in *OALD6*. To conclude, here are two points of interest. First, the taboo sign is attached to *cretin* in *OALD3* only. No other dictionary applies the taboo sign to this word, neither to *moron*, *idiot* or *imbecile*. Second, the taboo sign is found in the entries of all three words related to the taboos of body and sex, except in the *OALD5* entry for *dickhead*.

With one exception, unmodified labels are preferred in this semantic field. The only entry marked as '(+)' specifies an alternative context of use, namely *infml or derog* for *half-wit* in *OALD5*. This is yet another example of alternatives that may justly be criticized; the word *half-wit* does not surely lose its derogatory tone in informal contexts.

The set of labels assigned to *moron* serves as an example of the fact that the distinction between the labels *derogatory* and *offensive* is not clear. The word is described as follows: *derog* in *OALD4*, *derog or offensive* in *OALD5* and "an offensive word" in *OALD6*.

As to the consistency of labelling, thirty-six words for a stupid person appear unlabelled in all dictionaries listing the word. Conversely, five words are labelled in all dictionaries. Together, these forty-one words (76%) make up the group of uniform labelling. Near-uniformity occurs on five words. No uniformity is found in the labelling of eight words: *bonehead*, *booby*, *dolt*, *fool*, *imbecile*, *jerk*, *knucklehead* and *pinhead*.

The instances of using the taboo sign as a cautionary mark were already discussed above. Turning the instances into figures of consistency, twelve words can be taken into consideration, as only twelve words are labelled negatively in more than one dictionary. Eight words are consistently left without the taboo sign, whereas two words (*twat* and *wanker*) represent the opposite. The remaining two words are treated with more variation.

8.8 Treatment of Adjectives for Fat

Words in this semantic field are chosen to describe overall fatness in a person, not only fatness of a certain body part, such as fingers or face. The modern society appreciates thin and healthy looks to such an extent that even slight fatness is sometimes taken as a physical defect. Despite the fact, the number of overweight people is rising around the world. Considering this contradiction, words for fatness make up a taboo loaded area of vocabulary. In Ullmann's classification, it is a question of a taboo of delicacy.

8.8.1 Inclusion

The number of adjectives for 'fat' ranges between thirty-two of *OALD1* to forty-three of *OALD6*, as stated in Table 8b. *OALD2* adds seven items to its coverage, going up from the original thirty-two to thirty-nine. That was the largest addition in the line of dictionaries. Quite interestingly, the number remains on the same level in following editions: *OALD3, OALD4* and *OALD5* all list forty-two equivalents for *fat*. *OALD6* covers yet one term more, listing forty-three words of the forty-six under investigation.

Two words (*porky* and *heavyset*) were ignored as to consistency of inclusion, because they appear in *OALD6* for the first time. Thirty-one of the forty-six words are listed in all editions and another ten remain listed after they have entered the dictionary. Put together, these groups form 89% of the semantic field. Two items (*broad in the beam* and *fattish*) fluctuate; in other words, their status changes more than once. *Pursy* is the only word to disappear from the dictionaries.

8.8.2 Labelling

Words describing fatness are labelled negatively in *OALD4* for the first time, even though this semantic field is well covered in the earlier editions, too (Table 8a). The terms used as labels are *derog* (*OALD4* and *OALD5*) and *disapproving* (*OALD6*). Further, *OALD4* and *OALD6* refer the reader to a separate usage note, in which a number of adjectives for 'fat' are discussed. The descriptions "not polite", "more insulting than", "less offensive than" and "can sound offensive" are found in these notes (Table 8a).

The frequency of cautionary labels (Table 8b) remains quite stable in *OALD4* and the successive editions: 12%-19% of listed words. The taboo sign is not attached to any of the items in the group. *OALD4* and *OALD5* show a slight preference for unmodified labels, whereas in *OALD6*, the numbers are equal. However, as the number of labelled words is rather low, a comparison of the numbers of unmodified labels against modified labels is not very reliable. The adverbial that modifies the label is in all cases *usu* (for *usually*). Verbal descriptions that point out inoffensive situations of use are also treated as cases of modified labelling. Of the descriptions listed above, "less offensive than" and "can sound offensive" are taken as implying a possibility for an inoffensive use.

An inoffensive situation of use is specified on one adjective in this semantic field; the word *elephantine* is labelled *derog or joc* in *OALD4* and *OALD5*. In other words, if not used in a derogatory tone, the situation of speech is likely to be humorous. One might still ask whether there is a possibility that the referent feels offended, even though the aim is to be humorous in a positive sense. Quite curiously, *OALD6* labels the word *elephantine* as *formal* or *humorous*, without any cautionary label.

Multiple cautionary markings are found only when a label is complemented by a separate usage note. Sometimes the contents of a usage note contradict the label given in the entry. In such cases, the marking implying the strongest opposition overrules the other. For example, the word

podgy is labelled *usu derog* in the entry of *OALD4*, and is reported to be "more insulting than fat" in the usage note, receiving the mark '+'.

An examination of the entries of *podgy* and *pudgy* (in *OALD4* and *OALD5*) reveals that dictionary-making has not always been a very systematic business. *Podgy* is labelled *usu derog* and discussed in the usage note of both dictionaries. *Pudgy*, albeit clearly being a variant of the *podgy*, receives a cautionary mark in neither of the dictionaries. The *OALD6* entry for *pudgy* refers the reader to the entry of *podgy*, which is labelled *usu disapproving*.

Thirty-five of the sample words (80%) are unlabelled in all dictionaries that list the word. Quite interestingly, none of the words in this semantic field receive a unanimous cautionary labelling. *OALD6* assigns a cautionary label to four words (*heavy*, *large*, *pudgy* and *squat*), while the others do not. That equals 9%. Five words (*blowsy/blowzy*, *elephantine*, *fat*, *flabby* and *podgy*), equalling 11%, are labelled in more than one dictionary, but are also unlabelled in more than one. That is to say, they are labelled inconsistently.

9. Summary of the Findings

In this chapter, I will integrate the findings of Chapter 8 into a comparison of the six dictionaries of *OALD* against each other, as well as of the eight semantic fields against each other.

9.1.1 Inclusion

Table 9a shows percentages of word inclusion in each semantic field and each dictionary. The original figures are turned into percentages to allow easier comparison. When the eight word groups are put together, the overall percentages of words found in the dictionaries are as follows: 36% (*OALD1*), 53% (*OALD2*), 67% (*OALD3*), 82% (*OALD4*), 76% (*OALD5*) and 85% (*OALD6*). Interestingly enough, while all the other dictionaries exceed the number of words listed by the previous edition, *OALD5* lists fewer words of the sample material than its predecessor. This is a somewhat unexpected result, considering the fact that *OALD5* was the first edition based on a corpus. One would think that the introduction of a corpus would result in an increased number of controversial words.

The overall numbers of word inclusion quite naturally reflect the situation in several word groups. There are, however, differences between the semantic fields, as Table 9a indicates. For instance, the semantic field 'fat' receives remarkably high figures in every dictionary. Secondly, in contrast to what was said above about the overall inclusion rates of *OALD5*, in the group 'having sex', *OALD5* exceeds the percentage of *OALD4* by two percent. In the semantic fields 'homosexual' and 'exclamations of anger', the figures remain the same.

Further, some groups tend to be better covered than others, and the groups change over time. In other words, in the earlier editions, certain word groups ('homosexual', 'having sex', 'buttocks' and 'exclamations') receive percentages that are lower than the overall percentage of those editions. In the newer editions, word groups with relatively low percentages are 'prostitute', 'ethnic minority person' and 'stupid person'. With the exception of 'prostitute', this result correlates with the distinction between taboos of propriety and taboos of delicacy. On this point, the results of

this study differ from those of Wachal's study (see Chapter 3.2 on page 18): Wachal observed that offensive words referring to human body are often missing from current dictionaries, and conversely, that ethnic terms are relatively well covered nowadays.

The consistency in which words are kept listed in the dictionaries is summarized in Table 9b. Of the whole sample material (272 expressions), 78 are listed in all six dictionaries. That equals 29%. Another 44% remain listed once they have been listed for the first time. 4% of the items have been listed in the first edition and possibly also in some of the following editions, but not in the last two editions. They are considered as having disappeared from the vocabulary as described in learner dictionaries. 13% of the items fluctuate: their status changes more than once in the line of dictionaries. *OALD6* made a change into the inclusion status of 10% of the sample material: it added twenty-five new words and omitted three words. Whether these are temporary or more stable changes remain to be seen, therefore these are placed into a residue group here.

When the semantic groups are regarded separately, it can be seen that the groups 'ethnic minority' and 'fat' are early developers: a considerable proportion of the expressions in those groups have been included in all six dictionaries. The categories of 'words listed in all editions' and 'words which remain' put together form the most stable part of each semantic field. The semantic fields in which that figure exceeds the average of 73% (29% + 44% of the whole sample material) are 'homosexual', 'buttocks', 'exclamations of anger' and 'fat'. Thus, in those semantic fields, words are kept listed more often than in the other fields. For the group of 'exclamations of anger', this result is not very reliable, as the original set of search words was gathered using the *OALD6*. The proportion of words that remain listed is likely to be higher than it would be, if the sample material was more balanced in a temporal perspective. For the same reason, no items seem to have disappeared from the group of exclamations.

Quite interestingly, the word group 'ethnic minority person' seems to have lost a considerably greater proportion of words than the other groups (12% against the average of 4%). In

the group 'prostitute', instances of fluctuation are more usual than in the other groups (30% against the average of 13%). These results show no clear correlation with the classification of taboo types.

9.1.2 Labelling

Terms used as cautionary labels are not very numerous in the series of dictionaries. *OALD1* uses the terms *contemptuous* and *derogatory* as cautionary labels. The attributes (written in the full form here) *contemptuous*, *derogatory*, *impolite*, *not in polite use*, *term of abuse*, *vulgar* and the taboo sign ⚠ are found in *OALD2*. *OALD3* uses the same set as *OALD2* with the addition of *offensive*. This is quite a surprising finding, as the list of labels in the user guide does not include *contemptuous*, *impolite*, *not in polite use* or *offensive*. In contrast, the term *pejorative* is found on the list, but does not appear in the sample material. The labels found in *OALD4* are *derogative*, *offensive*, *sexist* and the taboo sign ⚠, which coincide with the labels of the user guide. The user guide of *OALD5* offers the same set of cautionary labels, and all but *sexist* are found in the sample material. *OALD6* uses the labels *disapproving*, *offensive* and the taboo sign ⚠, all of which are included in the inventory of labels as set in the user guide. In addition, attributes such as *insulting*, *not polite*, *rude*, *very rude* are used in verbal usage notes found in the *OALD6*. This comparison and the evaluation of the user guides indicates that some attention was paid to the labelling policies in *OALD3*, but the work was not very systematic until in *OALD4*, in which several pages are given over to explaining aspects of usage marking. Since then, the sections on usage marking have been reduced, but at the same time, have become easier to follow.

The frequencies of cautionary labels in each semantic field and each dictionary are recapitulated in Table 9c. The overall percentages are indicated in Table 9d, from which it can be seen that *OALD1* assigns a cautionary label to 2% of all words that are listed in the dictionary. For the other dictionaries, the corresponding figures are 8% (*OALD2*), 19% (*OALD3*), 32% (*OALD4*), 30% (*OALD5*) and 31% (*OALD6*). In short, the earlier editions up to *OALD4* show a steady

increase in the use of labels. Since then, the frequencies have remained almost exactly at the same level.

Table 9c shows that there are differences in the labelling frequencies of the eight semantic fields. Words in the groups 'homosexual' and 'ethnic minority person' are assigned a cautionary label relatively often, whereas these labels are less often attached to words for 'buttocks' and 'fat'. This result coincides to some extent with Norri's lecture observation (see Chapter 4, footnote on p. 23) about labelling frequencies of different taboo types. Norri suggested that taboos of propriety are labelled more often than taboos of delicacy.

It can be seen from Table 9d that the dictionaries (except for *OALD1*) prefer unmodified labels to modified ones. Occasional exceptions are found in some semantic fields and some dictionaries. For example, in the semantic field 'prostitute', *OALD4* has two entries marked as '+' against four entries of '(+)'. In *OALD5*, the cases are one against four. The modifier that restricts the application of the label is most often *usu/usually*, but adverbials *sometimes* and *often* are also found. Although the distinction between these modifiers is not perhaps very clear to a foreign language learner, all three items serve the purpose quite well. Modifiers indicating a temporal condition to the label, such as *now* and *formerly*, are found in the sample material. Moreover, the modifier *esp* appears in a few entries. This modifier is perhaps not a very good one for a learner dictionary, as it is somewhat ambiguous.

Alternative, inoffensive situations of use are also indicated occasionally in the dictionaries. However, these alternatives are often of questionable nature, as the members do not exclude each other, as noted by Norri (see Chapter 3.1, p. 16). An example of such alternatives is found in the entry of *half-wit* in *OALD5*, as it is labelled *infml or derog*. Alternatives that are more acceptable are found, too: for instance, the entry of *queer* in *OALD6* combines the taboo sign to a verbal note: 'also used by some homosexuals about themselves'.

No calculations were made on the use of multiple cautionary labels, but it can be noted that strings of more than one cautionary label are quite frequent in the sample material. Most often these

are combinations of the taboo sign and another cautionary label. A combination of *derog* (for *derogatory*) and *offensive* are also found. Norri noted in his study (2000: 77) that the distinction between *derogatory* and *offensive* made in the user guides of some dictionaries does not hold in practice (for further discussion, see Chapter 3.1, p. 15), and the choice between the two often appears haphazard. Considering the sample material of this study, the taboo sign could be added to the list. For example, the descriptions of *pansy* do not show any clear trend: ⚠ *derog, offensive term* (*OALD3*), *derog* (*OALD4* and *OALD5*) and ⚠ *offensive word* (*OALD6*).

The figures for consistency of labelling for the entire sample are found in the bottom row of Table 9e. 76% of the words that can be analysed on this matter are labelled similarly in all dictionaries listing the word. In other words, 76% of words listed in more than one dictionary appear either unlabelled in all dictionaries or labelled in all dictionaries. For 26%, one dictionary acts differently from the others. The labelling of 13% varies even more.

When the semantic fields are compared, the groups that are labelled more harmoniously than the sample material as a whole are 'buttocks', 'exclamations of anger', 'homosexual', 'having sex' and 'fat'. The semantic field 'stupid person' represents the whole sample material quite well, as its percentages are quite close to those of the entire sample material. In contrast, semantic fields with relatively high figures of non-uniformity are 'ethnic minority' and 'prostitute'. Quite interestingly, the results do not match with those of Norri (2000). Whereas words for racial and cultural background were labelled harmoniously in the material of Norri's study, the group 'ethnic minority person' received the highest figures of inconsistency in this study. A possible reason is the diachronic perspective of this thesis. There may have been a semantic change going on, which could not be seen in Norri's contemporary material. Conversely, words for sexual orientation and physical appearance were labelled inconsistently in Norri's material, whereas in the current sample material, the groups 'homosexual' and 'fat' were treated quite harmoniously.

Table 9f and Table 9g sum up the information on the use of the taboo sign ⚠. The overall frequencies are as follows: 0% (*OALD1*), 45% (*OALD2*), 76% (*OALD3*), 40% (*OALD4*), 37%

(*OALD5*) and 54% (*OALD6*). In other words, *OALD1* does not use the taboo sign at all, whereas in *OALD3*, the taboo sign is found in 76% percent of dictionary entries receiving a cautionary label. Whereas *OALD2*, *OALD3* and *OALD6* show a preference for combining the taboo sign with another cautionary label or note, *OALD4* and *OALD5* opt more often for using the taboo sign as a single label.

When the eight semantic fields are considered separately, the numbers of negatively labelled entries are quite low and thus not quite reliable. Therefore, a table recapitulating the frequencies of each semantic field is not included. The same applies, of course, to analysing the consistency in which the taboo sign is applied to individual items: the semantic fields are not compared against each other.

Considering the use of the taboo sign in the whole sample material, sixty-nine items can be taken into account. In other words, sixty-nine items of the whole sample material are labelled negatively in more than one dictionary. Fifty-one of these (equalling 74%) are treated uniformly, either the taboo sign appears in the cautionary label in all dictionaries or none of the dictionaries use the taboo sign as a cautionary label on a given word. On ten words (14%), one dictionary deviates from the practice of others, and finally, more variation is found in the cautionary labels of eight words (12%).

10. Concluding Remarks

In this thesis, I have studied inclusion and cautionary labelling of controversial topics and relating vocabulary in the *OALD* series of dictionaries. The sample material consisted of 272 expressions representing eight topics. Each of these topics were considered to form a semantic field. The semantic fields were first investigated separately. In the final section, a comparison was made between the six dictionaries as well as between the eight semantic groups.

It was found out that, overall, the trend is toward a more systematic description of the language. In other words, the modern dictionaries cover a larger part of the sample material as the older ones. Similarly, the approach to cautionary labelling has become more systematic. The most considerable step was taken in *OALD4* (published in 1987) when the dictionary provided the reader with a thorough guide into usage labelling. The frequency of cautionary labels has remained on the same level since then.

However, the eight semantic fields are not treated alike. Some semantic fields are covered more rigorously than the average of all six editions. Furthermore, some semantic fields receive relatively high percentages of inclusion in the older editions, but considerably lower in the current ones. The opposite is true for some other semantic fields. Similarly, the labelling practices vary between the semantic fields. The items of some semantic fields are labelled more often than items of other fields.

The variation in the treatment of the eight semantic groups could not be explained unequivocally. The classification of taboos into those of propriety and of delicacy is apparently not the only factor affecting the treatment of words. As Norri suggested, the definition of a word may reveal the negative tone, in which case a label may be considered redundant (e.g. 'fat'). Further, the topics may vary in their degree of controversy. For instance, words for 'buttocks' could perhaps be said to involve a milder type of controversy or unease, in which case cautionary markings are not necessary.

The most important shortcoming of this work was the research task that was probably too large and detailed for this kind of project. A more focused task would perhaps have brought depth into the analysis of each semantic field.

Conversely, the sample material was too narrow in some semantic fields, which made the quantitative analysis unreliable. For the same reason, this study was not able to bring out information about the use of the taboo sign as well as the intended.

Labels that describe other types of connotations than negative ones could be taken up as an area of further investigation. There are topics that require consideration as to the choice of words, but do not quite fit into the definition of controversial topics. An example of such a topic is dying. It would be worthwhile to study how dictionaries guide the reader in choosing between the numerous euphemistic expressions – some of which are quite humorous, while the others have a very serious overtone.

Bibliography

A. Primary Sources

Dictionaries

Cowie, A. P. (ed.) 1989. *Oxford Advanced Learner's Dictionary of Current English*. 4th ed. 1st ed. 1948. Oxford: Oxford University Press. (*OALD4*)

Crowther, Jonathan (ed.) 1995. *Oxford Advanced Learner's Dictionary of Current English*. 5th ed. 3rd impression. 1st ed. 1948. Oxford: Oxford University Press. (*OALD5*)

Hornby, A. S., E. V. Gatenby and H. Wakefield. (eds.) 1948. *A Learner's Dictionary of Current English*. London: Oxford University Press and Geoffrey Cumberlege. (Originally published under the title *Idiomatic and Syntactic English Dictionary*. 1942. Tokyo: Institute for Research in English Teaching.) (*OALD1*)

Hornby, A. S., E. V. Gatenby and H. Wakefield. (eds.) 1963 (1967). *The Advanced Learner's Dictionary of Current English*. 2nd ed. 7th impression. 1st ed. 1948. London: Oxford University Press. (*OALD2*)

Hornby, A. S., with A. P. Cowie and A. C. Gimson (eds.) 1974 (1983). *Oxford Advanced Learner's Dictionary of Current English*. 3rd ed. 17th impression. (11th revised and reset impression 1980.) 1st ed. 1948. Oxford: Oxford University Press. (*OALD3*)

Wehmeier, Sally (ed.) 2000 (2001). *Oxford Advanced Learner's Dictionary of Current English*. 6th ed. 6th impression. 1st ed. 1948. Oxford: Oxford University Press. (*OALD6*)

Wehmeier, Sally (ed.) 2000. *Oxford Advanced Learner's Dictionary of Current English, Electronic edition on CD-ROM*. 6th ed. 1st ed. 1948. Oxford: Oxford University Press. (Electronic edition first published 1997. Software: Copenhagen: TEXTware A/S.) (*OALD6 CD-ROM*)

Thesauri

Brown, Ivor (ed.) 1962 (1966). *Roget's International Thesaurus*. 3rd ed. 1st ed. n.d. London: Collins.

Chapman, Robert L. (ed.) 1992 (1996). *Collins Roget's International Thesaurus*. 5th ed. 1st. ed. 1852. Glasgow: HarperCollins.

Dutch, Robert A. (ed.) 1962 (1977). *Roget's Thesaurus of English Words and Phrases*. New revised and modernised ed. 1st ed. 1852. London: Longman.

Ferguson, Rosalind, Martin Manser and David Pickering (eds.) 2000. *The New Penguin Thesaurus*. London: Penguin Books.

Kirkpatrick, Betty (ed.) 1987. *Roget's Thesaurus of English Words and Phrases*. New ed. 1st ed. 1852. London: Longman.

Mawson, C. O. Sylvester (ed.) 1936 (1940). *Roget's Thesaurus of the English Language in Dictionary Form.* Revised ed. 1st ed. 1931. Garden City: Garden City Publishing Company.

Urdang, Laurence (ed.) 1991 (1997-1999). *The Oxford Thesaurus on CD-ROM.* Oxford: Oxford University Press. (Software: Copenhagen: TEXTware A/S and Focus Multimedia.)

B. Works Cited

Dictionaries

The Oxford English Dictionary Online. [Internet] Oxford: Oxford University Press. Available from < http://dictionary.oed.com/>. [Accessed 5 May 2008]

Other References

Allan, Keith and Kate Burridge 1991. *Euphemism and Dysphemism: Language Used as Shield and Weapon.* New York: Oxford University Press.

Allen, Robert 1996. "The Big Four: The Year of the Dictionaries." *English Today* 12:2, 41-51.

Burchfield, Robert 1989 (1992). "The Treatment of Controversial Vocabulary in the *Oxford English Dictionary*" in Robert Burchfield (ed.), *Unlocking the English Language.* London: Faber and Faber, 83-108. The article is a corrected reissue of an essay published in *Transactions of the Philological Society* 1973, 1-28.

Burkhanov, Igor 2003. "Pragmatic Specifications: Usage Indications, Labels, Examples; Dictionaries of Style, Dictionaries of Collocations" in Piet van Sterkenburg (ed.), *A Practical Guide to Lexicography.* Amsterdam: John Benjamins, 102-113.

Claire, Elizabeth 1998. *Dangerous English 2000! An Indispensable Guide for Language Learners and Others.* 3rd ed. 1st ed 1980. McHenry, Il: Delta Systems.

Cowie, A. P. 1998a. "A. S. Hornby, 1898-1998: A Centenary Tribute." *International Journal of Lexicography* 11:4, 251-268.

Cowie, A. P. 1998b. "A. S. Hornby, 1898-1998: Commemorative Issue: Introduction." *International Journal of Lexicography* 11:4, 249-250.

Delbridge, Arthur 1987. "Cultural, Situational and Modal Labels in Dictionaries of English" in Ross Steele and Terry Threadgold (eds.), *Language Topics: Essays in Honour of Michael Halliday*, vol. 2. Amsterdam: John Benjamins, 65-69.

Gramley, Stephan and Kurt-Michael Pätzold 1992 (1996). *A Survey of Modern English.* London: Routledge.

Hartmann, R. R. K. and Gregory James (eds.) 1998. *Dictionary of Lexicography.* London: Routledge.

Herbst, Thomas 1996. "On the Way to the Perfect Learners' Dictionary: A First Comparison of OALD5, LDOCE3, COBUILD2 and CIDE." *International Journal of Lexicography* 9:4, 321-357.

Hughes, Geoffrey 1988. *Words in Time: A Social History of the English Vocabulary.* Oxford: Blackwell.

Ilson, Robert F. 1986. "Lexicographic Archaeology: Comparing Dictionaries of the Same Family" in R. R. K. Hartmann (ed.), *The History of Lexicography*. Studies in the History of the Language Sciences, vol. 40. Amsterdam: John Benjamins, 127-136.

Landau, Sydney I. 2001. *Dictionaries: The Art and Craft of Lexicography*. 2nd ed. 1st ed. 1984. Cambridge: Cambridge University Press.

Leach, Edmund 1964. "Anthropological Aspects of Language: Animal Categories and Verbal Abuse" in Eric H. Lenneberg (ed.), *New Directions in the Study of Language*. Cambridge, MA: MIT Press, 23-63.

Lipka, Leonhard 1990. *An Outline of English Lexicology: Lexical Structure, Word Semantics, and Word-Formation*. Tübingen: Niemeyer.

Ljung, Magnus 1984. *Om svordomar i svenskan, engelskan och arton andra språk*. Stockholm: Akademilitteratur.

Norri, Juhani 1999. "Insulting Words, Taboo Words, Euphemisms, and Archaic/Historical Words." Lecture on Lexicology at the University of Tampere, 18 October 1999.

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

Palmer, F. R. 1981. *Semantics*. 2nd ed. 1st ed. 1976. Cambridge: Cambridge University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (eds.) 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Read, Allen Walker 1934. "An Obscenity Symbol." *American Speech* 9:4, 264-278. [Internet] Available from <http://links.jstor.org/sici?sici=00031283%28193412%299%3A4%3C264%3AAOS%3E2.0.CO%3B2-R>. [Accessed 17 February 2004]

Rothwell, J. Dan 1982. *Telling It Like It Isn't: Language Misuse & Malpractice / What We Can Do about It*. Englewood Cliffs, NJ: Prentice-Hall.

Rundell, Michael 1998. "Recent Trends in English Pedagogical Lexicography." *International Journal of Lexicography* 11:4, 315-342.

Strevens, Peter and Tom McArthur 1992. "Learner's Dictionary" in Tom McArthur and Feri McArthur (eds.), *The Oxford Companion to the English Language*. Oxford: Oxford University Press, 593-594.

Swan, Michael 1980 (1986). *Practical English Usage*. Oxford: Oxford University Press.

Svensén, Bo 1987. *Handbok i lexikografi: Principer och metoder i ordboksarbetet*. Stockholm: Esselte Studium and Tekniska nomenklaturcentralen.

Todd, Loreto and Ian Hancock 1986. *International English Usage*. London: Croom Helm.

Trask, R. L. 1995 (1996). *Language: The Basics*. London: Routledge.

Ullmann, Stephen 1962 (1972). *Semantics: An Introduction to the Science of Meaning*. Oxford: Blackwell.

Verkuyl, Henk, Maarten Janssen and Frank Jansen 2003. "The Codification of Usage by Labels" in Piet van Sterkenburg (ed.), *A Practical Guide to Lexicography*. Amsterdam: John Benjamins, 297-311.

Wachal, Robert S. 2002. "Taboo or not Taboo: That is the Question." *American Speech* 77:2, 195-206.

Appendix A: Analysis in a Tabular Form

1. Words for a prostitute

Table 1a. Labelling of words for a prostitute

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
call-girl	0	–	–	–	–	–
courtesan	–	–[1]	–[2]	–[3]	–[3]	–[4]
gigolo	0	0	0	+[5]	(+)[6]	–
harlot	–	(+)[7]	(+)[7]	(+)[8]	+[9]	+[10]
hooker	0	0	0	–	–	–
hustler	0	0	–	–	–	–
moll	0	–	–	0	0	0
prostitute	–	–	–	–	–	–
rent boy	0	0	0	0	0	–
scarlet woman	–	0	–[11]	(+)[12]	(+)[12]	0[13]
scrubber	0	0	0	+[5]	0	+[14]
streetgirl	0	–	–	–	0	0
streetwalker	–	–	–	–	–[15]	–[16]
strumpet	–	–[17]	–[17]	(+)[18]	(+)[18]	0
tart	0	–	+[5]	–	–	–
trollop	0	0	–	0	0	0
white slave	–	–	–	–	0	0
whore	–	+[19]	+[20]	(+)[21]	(+)[21]	–[16]
woman of easy virtue	0	0	0	–[22]	0	0
working girl	0	0	0	0	0	–[23]

[1] Labelled *in former times*.
[2] Labelled *in former times, esp in court circles.*
[3] Labelled *formerly*.
[4] Labelled *in the past*.
[5] Labelled *derog*.
[6] Labelled *usu derog*.

[7] Labelled *archaic, or as a term of abuse*.
[8] Labelled *arch or derog*.
[9] Labelled *arch derog*.
[10] Labelled *old use, disapproving*.
[11] Labelled *old use*.
[12] Labelled *dated derog or joc*.
[13] The sense 'prostitute' of *scarlet woman* is dropped out of *OALD6*, which now gives the sense 'woman who has sexual relationships with many different people' only, labelled *old-fashioned*.
[14] Described as "an offensive word…."
[15] Labelled *dated*.
[16] Labelled *old-fashioned*.
[17] Labelled *archaic*.
[18] Labelled *arch or joc derog*.
[19] Labelled *contemptuous term*.
[20] Labelled △ *derog*.
[21] Labelled *dated or derog*.
[22] Labelled *euph*.
[23] An explanatory note within the definition indicates a euphemistic use: "People say 'working girl' to avoid saying 'prostitute'." Also labelled *becoming old-fashioned*.

Table 1b. Frequency of labelling of words for a prostitute

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words included (20)	8	11	14	16	12	13
Words labelled +	0	1	2	2	1	2
Words labelled (+)	0	1	1	4	4	0
Percentage of labelling	0	18	21	38	42	15

Table 1c. Frequency of the taboo sign △ of words for a prostitute

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words labelled negatively	0	2	3	6	5	2
Words labelled △ only	0	0	0	0	0	0

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words labelled ⚠ with another label	0	0	1	0	0	0
Percentage of the taboo sign ⚠	–	0	33	0	0	0

2.   Words for a homosexual person

Table 2a. Labelling of words for a homosexual person

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
closet queen	0	0	0	–[1]	0[2]	0[3]
confirmed bachelor	0	0	0	0	0	–[4]
dike, dyke	0	0	+[5]	+[6]	(+)[7]	(+)[8]
fag, faggot	0	0	+[5]	+[9]	+[10]	+[11]
fairy	0	0	+[5]	+[9]	+[9]	+[12]
gay	0	0	–	–	–	–
homosexual	0	–	–	–	–	–
lesbian	0	–	–	–	–	–
nancy	0	–	0	0	0	0
pansy	0	–	+[13]	+[9]	+[9]	+[11]
ponce	0	0	0	0	+[14]	+[15]
poof, poofter	0	0	+[5]	+[9]	+[16]	+[11]
queen	0	0	+[5]	+[9]	+[9]	+[11]
queer	0	–	+[5]	+[9]	+[17]	(+)[18]

---

[1] The word appears unlabelled in an example sentence, s.v. *closet* (adj): "*I never knew he was a closet queen*, ie homosexual."

[2] The word *closet homosexual* appears unlabelled in an example sentence, s.v. *closet* (adj).

[3] The word *closet gay* unlabelled in an example sentence, s.v. *closet* (adj).

[4] The term is given in the example sentences for both *confirmed* and *bachelor*, furnished with an indication of an implied meaning: s.v. *bachelor*: "*He was a **confirmed bachelor*** (= a person who intended never to marry; often used in newspapers to refer to a homosexual man)."

[5] Labelled ⚠ *derog*.
[6] Labelled ⚠.
[7] Labelled *usu offensive*.
[8] Labelled ⚠ *usually offensive*.
[9] Labelled *derog*.
[10] Labelled *dated offensive*.
[11] Labelled ⚠. Described as "an offensive word…."
[12] Labelled *disapproving*. Described as "an offensive word…."
[13] Labelled ⚠ *derog*, *offensive term*.
[14] Labelled *offensive*.
[15] Described as "an offensive word…."
[16] Labelled *derog or offensive*.
[17] Labelled *derog offensive*.
[18] Labelled ⚠. Described as "an offensive word … which is, however, also used by some homosexuals about themselves."

Table 2b. Frequency of labelling of words for a homosexual person

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words included (14)	0	5	10	11	11	12
Words labelled +	0	0	7	7	7	6
Words labelled (+)	0	0	0	0	1	2
Percentage of labelling	–	0	70	64	73	67

Table 2c. Frequency of the taboo sign ⚠ of words for a homosexual person

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words labelled negatively	0	0	7	7	8	8
Words labelled ⚠ only	0	0	0	1	0	0
Words labelled ⚠ with another label	0	0	7	0	0	6

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Percentage of the taboo sign ⚠	–	–	100	14	0	75

3. Words for verbs for having sex

Table 3a. Labelling of verbs for having sex

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
ball	0	0	0	0	0	+[1]
bang	0	0	0	–[2]	0	+[1]
(be) intimate[3]	0	0	0	–[4]	–[4]	–[5]
bed	0	0	0	–	–	–[6]
bonk	0	0	0	0	–[7]	–
(commit) adultery[8]	–	–	–	–	–	–
consummate (a marriage)[9]	0	–	–	–[10]	–[10]	–[11]
copulate	0	–	–	–[10]	–[10]	–[12]
couple	0	0	0	–[13]	–[14]	–[11]
do (it)	0	0	0	0	0	–
fornicate	0	0	–	(+)[15]	(+)[16]	+[17]
fuck	0	0	+[1]	+[1]	+[1]	+[1]
get it on	0	0	0	0	0	–
get off	0	0	–	–	–	–
get your leg over	0	0	0	0	0	–
go all the way	0	0	0	0	0	–
go to bed	0	0	0	–	–	–
go with	0	0	0	–	–	–[6]

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
have	0	0	0	+[1]	−	−
(have) carnal knowledge[18]	0	0	0	0	−[19]	−[20]
(have) intercourse[21]	0	−	−	−[10]	−[10]	−[22]
have it away	0	0	+[1]	+[1]	−	−
have it off	0	0	+[1]	+[1]	−	−
(have) sex[23]	0	−	−	−	−	−
(have) sexual intercourse[24]	0	−	−	−	−[10]	−[11]
(have) sexual relations[25]	0	0	0	−	−	−[11]
hump	0	0	0	+[1]	+[1]	+[1]
knock up	0	0	+[26]	0	0	0
lay (pass. *get laid*)	0	0	−	+[1]	+[1]	+[1]
lie with	−[27]	−[28]	−[29]	0	0	0
make	0	0	0	+[30]	0	0
make it	0	0	0	−	−	−
make love	0	0	−	−	−	−
make out	0	0	0	0	−	−
poke	0	0	0	0	0	+[1]
roger	0	0	0	+[31]	+[1]	+[1]
score	0	0	0	−	−	−
screw	0	0	+[1]	+[1]	+[1]	+[1]
shag	0	0	0	+[1]	+[1]	+[1]
sleep with	0	−[32]	−[33]	−[4]	−[4]	−
spend the night	0	0	0	−[4]	−[4]	−
take	0	0	0	−	−	0

[1] Labelled ⚠.

[2] Listed as a phrasal verb *bang away* in the sense 'have vigorous sexual intercourse', labelled *sl*.
[3] The meaning was found under the adjective *intimate*.
[4] Labelled *euph*.
[5] Labelled *formal or law*.
[6] Labelled *old-fashioned*.
[7] Labelled *joc*.
[8] The meaning was found under *adultery*.
[9] The meaning was found under *consummate*.
[10] Labelled *fml*.
[11] Labelled *formal*.
[12] Labelled *technical*.
[13] Labelled *arch or rhet.*
[14] Labelled *rhet or euph.*
[15] Labelled *fml esp derog.*
[16] Labelled *fml usu derog.*
[17] Labelled *formal*, *disapproving*.
[18] The meaning was found under *carnal knowledge*.
[19] Labelled *law*.
[20] Labelled *old-fashioned or law*.
[21] The meaning was found under *intercourse*.
[22] The phrase *have intercourse* is labelled *formal* (s.v. *intercourse*), but under *sexual intercourse* the label for *have intercourse* is *informal*.
[23] The meaning was found under *sex*.
[24] The meaning was found under *sexual intercourse*.
[25] The meaning was found under *(sexual) relations*.
[26] *OALD3* is the only dictionary to define *knock up* as 'have sexual intercourse; make pregnant', labelled △ *vulg*. The two earlier editions do not list the verb at all, and the latest editions of *OALD* give the latter sense 'make pregnant' only, labelled △ in *OALD4* and *OALD5*; unlabelled in *OALD6*.
[27] Labelled *old use*.
[28] Labelled *old use*, *biblical*.
[29] Labelled *old use, biblical, now usu **sleep with***.
[30] Labelled *sexist*.
[31] Labelled △ *euph*.
[32] Labelled *euphemism*.
[33] Labelled *euphem*.

Table 3b. Frequency of labelling of verbs for having sex

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words included (42)	2	8	17	31	32	38
Words labelled +	0	0	5	10	6	10
Words labelled (+)	0	0	0	1	1	0
Percentage of labelling	0	0	29	35	22	26

Table 3c. Frequency of the taboo sign ⚠ of verbs for having sex

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words labelled negatively	0	0	5	11	7	10
Words labelled ⚠ only	0	0	4	9	6	9
Words labelled ⚠ with another label	0	0	1	0	0	0
Percentage of the taboo sign ⚠	–	–	100	82	86	90

4. Words for buttocks

Table 4a. Labelling of words for buttocks

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
arse	0	+[1]	+[1]	+[2]	+[2]	+[2]
ass	0	0	+[3]	+[2]	+[2]	+[2]
backside	0	+[4]	–	–	–	–
behind	–	–	–	–[5]	–[5]	–[6]
bottom	–	–	–	–	–	–
bum	0	–	–	–	–	–
bun, buns	0	0	0	0	0	–
butt	0	0	0	–	–	–
buttock, buttocks	–	–	–	–	–	–
cheek, cheeks	0	0	0	–	–	–
duff	0	0	0	0	0	–
fanny	0	0	–	–[7]	–[7]	–[7]

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
posterior, posteriors	0	–[8]	–[8]	–[9]	–[9]	–[10]
rear	0	0	0	–[11]	–[11]	–
rear end	0	0	0	0	–[11]	–
rump	–	–[12]	–[13]	–[9]	–[9]	–[10]
seat	–	–	–	–[14]	–[14]	–[15]
stern	0	0	0	–[16]	0	0
tail	0	0	0	–[17]	0	0

[1] Labelled ⚠, *not in polite use*.
[2] Labelled ⚠.
[3] Labelled *vulg*.
[4] Labelled *not used in polite society*.
[5] Labelled *euph*.
[6] An explanatory note within the definition: "People often say 'behind' to avoid saying 'bottom'."
[7] From *OALD4* on, the entry for the word *fanny* consists of two senses, the first one being 'female sexual organs', labelled as *British English* ⚠ *slang*. The sense 'buttocks' is used labelled *slang, especially American English*.
[8] Labelled *hum*.
[9] Labelled *joc*.
[10] Labelled *humorous*.
[11] Labelled *euph*.
[12] Labelled *jocularly, of a human being*.
[13] Labelled *joc, of a human being*.
[14] Labelled *fml*.
[15] Labelled *especially formal*.
[16] Labelled *esp joc*.
[17] Labelled *dated*.

Table 4b. Frequency of labelling of words for buttocks

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words included (19)	5	9	11	16	15	17
Words labelled +	0	2	2	2	2	2
Words labelled (+)	0	0	0	0	0	0
Percentage of labelling	0	22	18	13	13	12

Table 4c. Frequency of the taboo sign ⚠ of words for buttocks

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words labelled negatively	0	2	2	2	2	2
Words labelled ⚠ only	0	0	0	2	2	2
Words labelled ⚠ with another label	0	1	1	0	0	0
Percentage of the taboo sign ⚠	–	50	50	100	100	100

5. Words for exclamations of anger

Table 5a. Labelling of exclamations of anger

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
ballocks, bollocks	0	0	$+^1$	$+^2$	$+^2$	$+^2$
balls	0	0	$+^2$	$+^2$	$+^2$	$+^3$
blast (it)	0	–	–	–	–	–
blimey	0	$+^4$	$+^4$	–	–	–
blow (it)	0	–	–	–	–	$-^5$
bother (it)	–	–	–	–	–	–
boy	0	0	–	–	–	–
brother	0	0	0	–	–	$-^5$
bugger (it)	0	$+^6$	$+^2$	$+^2$	$+^2$	$+^2$
Christ	0	0	0	$+^2$	$+^2$	$+^7$
for Christ's sake	–	–	–	–	–	$(+)^8$
confound (it)	–	–	$-^9$	–	$-^9$	$-^5$

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
crikey	0	0	–	–	–	–[10]
damn (it)	–	–	–	–	–	–[11]
dang	0	0	0	0	0	–[12]
darn (it)	0	0	–	–[13]	–[13]	–[14]
dash (it)	0	–[15]	–[15]	–[16]	–[16]	–[5]
dear	–	–	–	–	–	–
doggone (it)	0	0	0	–	–	–
drat (it)	0	–	–[9]	–	–	–[5]
flip	0	0	0	–	–	0
fuck (it)	0	0	+[2]	+[2]	+[2]	+[7]
God	–	0	–	–	–	(+)[17]
Heavens	–	–	–	–	–	–[18]
heck	0	0	–[19]	–[13]	–[13]	–
hell	–	–	–	–	–	(+)[20]
Lord	–	–	–	–	–	–
man	0	0	0	–	–	–
shit	0	+[21]	+[22]	+[2]	(+)[23]	+[24]
shite	0	0	0	0	0	+[2]
shoot	0	0	0	0	0	–[25]
sod	0	+[26]	+[27]	+[27]	+[28]	+[29]
son of a bitch	0	0	0	+[30]	+[30]	+[31]
son of a gun	0	0	0	0	0	–[32]
sugar	0	0	0	0	0	–[25]

[1] Labelled △ *vulg.*
[2] Labelled △.
[3] Labelled △ with an explanatory note: "Less offensive ways to express this are '*Nonsense!*' or '*Come off it!*'."
[4] Labelled *vulg.*
[5] Labelled *old-fashioned.*
[6] Not listed as an interjection. The word has two nominal uses, "sodomite" and "a vulgar term of abuse", the latter of which is labelled *taboo.*
[7] Labelled △ and described as "a swear word that many people find offensive…."
[8] An explanatory note within the definition: "Some people find the use of *Christ*, *God* or *heaven* here offensive."
[9] Labelled *dated.*
[10] Labelled *old-fashioned, spoken.*

[11] The phrase *damn* is unlabelled, equivalent expressions given inside brackets: "**damn** (also *old-fashioned* **dammit**, **damn it**)."

[12] The word *dang* is unlabelled, but the definition explains its euphemistic use: "[A] mild swear word, used instead of *damn*."

[13] Labelled *euph*.

[14] The word *darn* is unlabelled as to attitude (*spoken*), but the definition explains its euphemistic use: "[U]sed as a mild swear word … to avoid saying 'damn'."

[15] The word *dash* is unlabelled, but the definition explains its euphemistic use: "[U]sed as a mild substitute for *Damn!*"

[16] Labelled *euph*. In addition, the definition explains its euphemistic use: "[U]sed as a milder way of saying *damn."*

[17] The exclamation *God* is unlabelled as to attitude (*spoken*), but an explanatory note within the definition indicates offensiveness: "Some people find this use offensive."

[18] Labelled *spoken*.

[19] Labelled *euphem*.

[20] The exclamation *hell* is unlabelled, but an explanatory note within the definition indicates offensiveness: "Its use is offensive to some people."

[21] Not listed as an interjection, but the noun *shit* is labelled ⚠ *not in polite use*.

[22] The interjection *shit* is labelled ⚠ *vulg*, while the nominal use is labelled ⚠ *not polite use*.

[23] Labelled *usu* ⚠.

[24] Labelled ⚠. Described as "a swear word that many people find offensive…." In addition, an explanatory note attached to the entry: "Less offensive exclamations to use are *blast*, *darn it* (especially *AmE*), *damn* or (*BrE*) *bother*."

[25] The word is unlabelled, but the definition explains its euphemistic use: "[U]sed … to avoid saying 'shit'."

[26] Not listed as an interjection. Its nominal use '(esp. as a vulgar term of abuse) sodomite' is labelled ⚠.

[27] Not listed as an interjection, but as an idiom *Sod (it)!* under the verb *sod*, labelled ⚠.

[28] Listed as a verb which is used in imperative only, labelled ⚠.

[29] Listed as a verb which is used in imperative only, labelled ⚠. In addition, described as "a swear word that many people find offensive…."

[30] Not listed as an interjection. As a noun, 'unpleasant person', it is labelled ⚠.

[31] Not listed as an interjection. As a noun, 'unpleasant person', it is labelled ⚠ and described as "an offensive word…."

[32] The phrase *son of a gun* is not listed as an interjection. As a noun it is unlabelled as to attitude (*spoken*). However, the example sentence indicates a use as an interjection: *"Well, son of a gun–and I thought the old guy couldn't dance!"*

Table 5b. Frequency of labelling of exclamations of anger

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words included (35)	9	16	24	30	30	34
Words labelled +	0	4	7	8	7	9
Words labelled (+)	0	0	0	0	1	3
Percentage of labelling	0	25	29	27	27	35

Table 5c. Frequency of the taboo sign △ of exclamations of anger

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words labelled negatively	0	4	7	8	8	12
Words labelled △ only	0	0	4	8	8	3
Words labelled △ with another label	0	3	2	0	0	6
Percentage of the taboo sign △	–	75	86	100	100	75

6. Words for a person belonging to an ethnic minority group

Table 6a. Labelling of words for a person belonging to an ethnic minority group

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
African American	0	0	0	0	–	–
Afro-American	0	0	–	–	0	0
American Indian	0	0	–	–	$-^1$	–
Amerindian	0	0	0	–	$-^1$	$-^2$
black, Black (noun)	–	–	$(+)^3$	$(+)^4$	$(+)^5$	$(+)^6$
black person, man, woman (i.e. *black* or *Black* as an attribute adjective)	0	0	0	–	$(+)^5$	$(+)^7$
blackamoor	–	$(+)^8$	$(+)^9$	$+^{10}$	0	0
boy	0	0	0	0	0	$+^{11}$

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
brave	0	–	–[12]	–	–	–[2]
buck	0	0	0	+[13]	0	0
coloured, Coloured (noun)[14]	0	0	0	0	+[15]	(+)[16]
coloured person, man, woman (i.e. *coloured* or *Coloured* as an attribute adjective)[14]	–	–	–	–[17]	(+)[18]	(+)[16]
coon	–	–	+[19]	+[19]	+[20]	+[21]
darkey, darkie, darky	0	–	+[22]	+[23]	0	0
half-breed	–	–	–	(+)[24]	(+)[25]	+[26]
half-caste	–	–	–	(+)[24]	(+)[25]	+[26]
Indian (noun)	–	–	–[27]	–	–	+[28]
Jim Crow	+[29]	+[29]	+[30]	+[31]	0	0
mammy	–	–	(+)[32]	(+)[33]	0	0
mulatto	–	–	–	–	–	+[15]
native, Native (noun)	(+)[34]	–	–	(+)[35]	(+)[36]	+[37]
Native American	0	0	0	0	–	–
Negress, negress	–	–	–	(+)[5]	(+)[5]	(+)[38]
Negro, negro	–	–	–	(+)[5]	(+)[5]	(+)[38]
Negroid, negroid (noun)	0	–	–	–	0	0
nigger	–	+[39]	+[40]	+[20]	+[20]	+[21]

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
non-white (noun)	0	0	0	–	–	–
person, man, woman of colour	–[41]	–	0	0	0	–
piccaninny, pickaninny	–	–	–[42]	+[43]	0	0
Red Indian	–	–	(+)[44]	+[23]	+[45]	+[46]
redskin	–[47]	–[47]	(+)[44]	+[23]	+[45]	+[48]
squaw	–	–	–	–	–	(+)[49]
uncle Tom	0	0	+[19]	+[13]	+[13]	+[50]
wog	0	0	0	0	+[23]	+[21]

[1] Labelled *dated.*

[2] Labelled *old-fashioned.*

[3] Labelled *formerly derog, but now widely used.*

[4] Labelled *formerly derog, now the preferred word.*

[5] Labelled *sometimes offensive.*

[6] An explanatory note within the definition: "In this meaning *black* is more common in the plural. It can sound offensive in the singular. Instead, you can use the adjective ('a black man/woman') or, in the US, *African American*."

[7] An explanatory note within the definition: "*Black* is the word most widely used and generally accepted in Britain. In the US the currently accepted term is *African American*. "

[8] Labelled *hum. or derog.*

[9] Labelled *old use, hum or derog.*

[10] Labelled *dated derog offensive.*

[11] Labelled △ and described as "an offensive way of addressing a black man."

[12] Labelled *poet.*

[13] Labelled *derog.*

[14] The focus here is on the term *coloured* in the sense 'person who does not have a white skin'. The South African use of *Coloured* (or *Cape Coloured*) 'person whose parents are of different races' is ignored here, even though it is listed in *OALD2-6* (unlabelled).

[15] Labelled *offensive.*

[16] Labelled *old-fashioned or offensive.*

[17] Labelled *becoming dated.*

[18] Labelled *dated often offensive.*

[19] Labelled △ *derog.*

[20] Labelled △ *derog offensive.*

[21] Labelled △ and described as "a very offensive word…."

[22] Labelled △ *offensive term.*

[23] Labelled △ *offensive.*

[24] Labelled *sometimes derog.*

[25] Labelled *usu offensive.*

[26] Labelled △ *offensive.* In addition, an explanatory note within the definition: "It is more acceptable to talk about 'a person *of mixed race*'."

[27] Listed as *American Indian* only; unlabelled.

[28] Labelled *old-fashioned*, *offensive.*

[29] Described as "a contemptuous name…."

[30] Labelled △ and described as "a contemptuous name…."

[31] Labelled *derog offensive.*

[32] Labelled △ *old use, now derog.*

---

[33] Labelled *dated now offensive.*

[34] An explanatory note within the definition: "In this sense the word is sometimes, but not always, derogatory. Whether it is derogatory depends upon the speaker and the situation."

[35] Labelled *esp offensive.*

[36] Labelled *dated usu offensive.*

[37] Labelled *old-fashioned, offensive.* In addition, a explanatory note within the definition: "[A] word used in the past…."

[38] Labelled *old-fashioned, often offensive.*

[39] Labelled △ *impolite word.*

[40] Labelled △ *impolite and offensive word.*

[41] The phrase *gentleman (lady) of colour* is labelled *colloq. joking style.*

[42] Labelled *old use.*

[43] Labelled △ *dated offensive.*

[44] Labelled *old use, now impolite.*

[45] Labelled △ *dated offensive.*

[46] Labelled *old-fashioned* △ and described as "a very offensive word…."

[47] Described as "an old name…."

[48] Labelled *old-fashioned* △ *offensive.*

[49] Labelled *old use* and described as "a word … that is now often considered offensive."

[50] Labelled △ *offensive* with an explanatory note: "[S]ometimes used in the past to refer to a black man…."

Table 6b. Frequency of labelling of words for a person belonging to an ethnic minority group

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words included (34)	19	22	24	28	24	26
Words labelled +	1	2	5	10	7	12
Words labelled (+)	1	1	5	7	8	7
Percentage of labelling	11	14	42	61	63	73

Table 6c. Frequency of the taboo sign △ of words for a person belonging to an ethnic minority group

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words labelled negatively	2	3	10	17	15	19
Words labelled △ only	0	0	0	0	0	0

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words labelled ⚠ *with another label*	0	1	6	6	5	9
Percentage of the taboo sign ⚠	0	33	60	35	33	47

7.   Words for a stupid person

Table 7a. Labelling of words for a stupid person

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
addlehead	0	–	–	0	0	0
airhead	0	0	0	0	0	$+^1$
ass	–	–	–	–	–	–
berk	0	0	0	$+^2$	$+^2$	$+^1$
birdbrain	0	0	0	0	0	–
blockhead	–	–	–	–	–	–
bonehead	0	–	–	$+^2$	$+^2$	0
boob	0	–	–	0	0	–
booby	–	–	–	$+^3$	$+^3$	0
chump	0	–	–	–	$-^4$	$-^5$
clod	0	0	0	0	0	–
clot	0	0	$-^4$	$-^6$	$-^6$	$-^5$
cretin[7]	0	0	0	$+^8$	$+^9$	$+^{10}$
dickhead	0	0	0	0	$+^2$	$+^{11}$
dimwit	0	0	0	–	–	–
dingbat	0	0	0	0	0	–
dolt	–	–	–	$+^2$	$+^2$	$+^1$
donkey	–	$-^{12}$	$-^{12}$	–	0	0

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
dope	0	0	0	−	−	−
dullard	−	−	−	−	0	−[5]
dumbbell	0	0	0	−	−	−
dummy	0	0	0	−	−	−
dunce	−	−	−	−	−	−[5]
dunderhead	0	−	−	+[2]	0	0
fat-head	−	−	−	−	−	0
fool	−	−	−	+[2]	+[2]	−
goof	0	0	−	−	−	−
goon	0	−	−	−	−	−[5]
goose	−	0	−	−[4]	−[4]	−[5]
half-wit	0	−	−	−	(+)[13]	−
idiot[7]	−	−	−	−	−	−
ignoramus	−	−	−	−	−	−[14]
imbecile[7]	−	−	−	−	+[2]	+[15]
jackass	−	−	−	−[16]	−	−
jerk	0	0	−	+[2]	+[2]	−
knuckle-head	0	0	0	+[2]	0	−
lunkhead	0	0	0	0	0	−
moron[7]	0	0	−	+[2]	+[17]	+[18]
muggins	0	0	−	−[6]	−[6]	−[19]
mutt	0	−	−	−	0	0
muttonhead	0	−	−	+[2]	0	0
nerd	0	0	0	0	0	+[1]
nincompoop	−	−	−	−	−[4]	−[5]
ninny	−	−	−	−	−	−
nitwit	0	−	−	−	−	−
numskull, numbskull	−	−	−	+[2]	0	−

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
nutter	0	0	0	–	–	–
oaf	0	0	0	–	–	–
pinhead	0	0	–	+[2]	0	0
rattlebrain	–	–	–	0	0	0
rattlehead	–	0	0	0	0	0
rattlepate	–	–	–	0	0	0
sap	0	–	–[4]	–	–	–
scatterbrain	–	–	–	–	–	–
silly	–	–	–	–	–	–
simpleton	–	–	–	–	–[4]	–[5]
thickhead	–	0	0	0	0	0
twat	0	0	0	+[20]	+[20]	+[21]
twerp	0	–	–	–	–	–[5]
twit	0	0	–	–[22]	–[22]	–
wally	0	0	0	–	–	–
wanker	0	0	0	+[20]	+[20]	+[23]

[1] Labelled *disapproving.*

[2] Labelled *derog.*

[3] Labelled *dated derog.*

[4] Labelled *dated.*

[5] Labelled *old-fashioned.*

[6] Labelled *joc.*

[7] *Moron* has two closely related senses, one used in medical contexts, and another which has developed from the first one through generalisation into the sense 'stupid person' appearing in general language use. The latter sense is regarded relevant here. Other similar words are *cretin*, *idiot* and *imbecile*.

[8] Labelled △ *offensive.*

[9] Labelled *offensive.*

[10] Labelled *spoken, offensive.*

[11] Labelled △ and described as "a very rude way of referring to sb, especially a man, that you think is stupid."

[12] The meaning is not explicitly listed as the entry for *donkey* only makes a reference to the entry for *ass.*

[13] Labelled *infml or derog.*

[14] Labelled *usually humorous.*

[15] Said to be "a rude way to describe a person that you think is very stupid."

[16] Labelled *fig. infml.*

[17] Labelled *derog or offensive.*

[18] Described as "an offensive way of referring to sb that you think is very stupid."

[19] Labelled *humorous.*

[20] Labelled △ *derog.*

[21] Labelled △ and described as "an offensive word for an unpleasant or stupid person."

[22] Labelled *often joc.*

[23] Labelled △ and described as "an offensive word used to insult sb, especially a man, and to show anger or dislike."

Table 7b. Frequency of labelling of words for a stupid person

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words included (62)	24	33	41	49	42	49
Words labelled +	0	0	0	15	12	10
Words labelled (+)	0	0	0	0	1	0
Percentage of labelling	0	0	0	31	31	20

Table 7c. Frequency of the taboo sign ⚠ of words for a stupid person

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Words labelled negatively	0	0	0	15	13	10
Words labelled ⚠ only	0	0	0	0	0	0
Words labelled ⚠ with another label	0	0	0	3	2	3
Percentage of the taboo sign ⚠	–	–	–	20	15	30

8.   Words for adjectives of fat

Table 8a. Labelling of adjectives for fat [1]

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
ample	–	–	–	–	–[2]	–
beefy	0	–	–	–	–	–

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
big	–	–	–	–	–	–
big-bellied	0	–	–	–	–	–
bloated	–	–	–	–	–	–
blowsy, blowzy	–	–	–	+[3]	+[3]	+[4]
broad	–	–	–	–	–	–
broad in the beam	0	0	–	–	–	0
bulbous	–	–	–	–	–	–[5]
bulky	–	–	–	–	–	–
buxom	–	–	–	–[6]	–[7]	–
chubby	–	–	–	–[8]	–[8]	–
chunky	0	0	–	–	–	–
corpulent	–	–	–	–[9]	–[9]	–[10]
dumpy	–	–	–	–	–	–
elephantine	0	–	–	(+)[11]	(+)[11]	–[12]
fat	–	–	–	–[8]/+[13]	–[8]/+[13]	–/+[13]
fattish	0	–	–	–	0	0
flabby	–	–	–	+[3] ([8])/+[14]	+[3] ([8])	+[4]/(+)[15]
fleshy	–	–	–	–	–	–
full	0	–	–	–	–[2]	–[16]
gross	–	–	–	–	–	–
heavy	–	–	–	–	–	–/(+)[17]
heavyset	0	0	0	0	0	–
large	–	–	–	–[2]	–[8]	–/(+)[17]
meaty	–	–	–	–	–[18]	–
obese	0	–	–	–[19] ([8])	–[19] ([8])	–[20]
overweight	–	–	–	–[8]	–[8]	–
paunchy	0	0	–	–	–	–

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
plump	–	–	–	–[21] ([8])	–[22] ([8])	–
podgy	–	–	–	(+)[23]([8])/ +[14]	(+)[23] ([8])	(+)[24]
porky	0	0	0	0	0	+[4]
portly	–	–	–	–	–	–
pot-bellied	0	–	–	–	–	–
pudgy	–	–	–	–	–	(+)[24]
pursy	–	–	–[25]	0	0	0
roly-poly[26]	0	0	0	0	–	–
rotund	–	–	–	–[27]	–[28]	–[29]
round	–	–	–	–	–	–
squat	–	–	–	(+)[23]	–	–
stocky	–	–	–	–	–	–
stout	–	–	–	–[30] ([8])	–[30] ([8])	–
thick	–	–	–	–	–	–
thickset	–	–	–	–	–	–
tubby	–	–	–	–[8]	–[8]	–
well-rounded	0	0	0	–	–	–

[1] A number of of the entries in *OALD4* and *OALD6* refer the reader to a separate usage note. Sometimes the contents of the note contradict the label given in the entry. In such cases, the marking implying the strongest opposition overrules the other. For example, the word *fat* appears unlabelled in the entry, but the usage note describes it as "not polite". *OALD5* has a similar note, too, but unlike the other two dictionaries, it lists ways in which the idea of fatness can be addressed in a positive tone.

[2] Labelled *euph*.

[3] Labelled *derog*.

[4] Labelled *disapproving*.

[5] Labelled *written*.

[6] Labelled *usu approv esp joc*.

[7] Labelled *usu approv*.

[8] The entry contains a cross-reference to a usage note under the entry for *fat*.

[9] Labelled *fml esp euph*.

[10] Labelled *formal*. In addition, an explanatory note within the definition: "People say 'corpulent' to avoid saying 'fat'."

[11] Labelled *derog or joc*.

[12] Labelled *formal or humorous*.

[13] Described as "not polite" in a usage note.

[14] Described as "more insulting than *fat*" in a usage note.

[15] Described as "can sound offensive " in a usage note.

[16] An explanatory note within the definition: "'Full' is sometimes used to avoid saying 'fat'."

[17] Described as "less offensive than *fat*" in a usage note.

[18] Labelled *approv*.

[19] Labelled *fml or medical*.
[20] Labelled *formal or medical*.
[21] The word is given two senses. The first one, 'having a full rounded shape; fleshy', is unlabelled. The second sense, 'overweight, fat', is labelled *euph*.
[22] The word is given two senses. The first one, 'having a pleasantly full round shape; fleshy', is labelled *usu approv*. The second sense, 'fat', is labelled *euph*.
[23] Labelled *usu derog*.
[24] Labelled *usually disapproving*.
[25] Labelled *old use*.
[26] In *OALD1-4*, the noun *roly-poly* is given two senses: 'dessert' and 'plump person', the latter appearing either unlabelled or labelled *colloq* or *infml*. In these dictionaries, the adjectival use is not listed. In *OALD5-6*, the noun is assigned only one meaning, 'dessert', and the word is also listed as an adjective 'short and fat', labelled *infml* and *informal*.
[27] Labelled *euph or joc*.
[28] Labelled *fml or joc*.
[29] Labelled *formal or humorous*.
[30] Labelled *esp euph*.

Table 8b. Frequency of labelling of adjectives for fat

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words included (46)	32	39	42	42	42	43
Words labelled +	0	0	0	4	3	4
Words labelled (+)	0	0	0	2	2	4
Percentage of labelling	0	0	0	14	12	19

Table 8c. Frequency of the taboo sign △ of adjectives for fat

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words labelled negatively	0	0	0	6	5	8
Words labelled △ only	0	0	0	0	0	0
Words labelled △ with another label	0	0	0	0	0	0

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Percentage of the taboo sign ⚠	–	–	–	0	0	0

9. Summary tables

Table 9a. Words listed in each semantic field in each dictionary (%)

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Prostitute (20)	40%	55%	70%	80%	60%	65%
Homosexual (noun) *(14)*	0%	36%	71%	79%	79%	86%
To have sex (42)	5%	19%	40%	74%	76%	90%
Buttocks (19)	26%	47%	58%	84%	79%	89%
Exclamation of anger (35)	26%	46%	69%	86%	86%	97%
Ethnic minority (34)	56%	65%	71%	82%	71%	76%
Stupid person (62)	39%	53%	66%	79%	68%	79%
Fat (adjective) *(46)*	70%	85%	91%	91%	91%	93%
Total (272)	36%	53%	67%	82%	76%	85%

Table 9b. Degree of uniformity in word inclusion in different semantic fields

	Words whose trend is unknown[1]		Words listed in all six editions		Words which remain listed once taken in		Words which disappear		Words which fluctuate[2]	
Prostitute (20)	2+1	15%	5	25%	5	25%	1	5%	6	30%
Homosexual (noun) (14)	1+0	7%	0		11	79%	0		2	14%
To have sex (42)	6+0	14%	1	2%	30	71%	1	2%	4	10%
Buttocks (19)	2+0	11%	5	26%	10	53%	0		2	11%
Exclamation of anger (35)	5+0	14%	8	23%	20	57%	0		2	6%
Ethnic minority (34)	1+0	3%	14	41%	10	29%	4	12%	5	15%
Stupid person (62)	6+2	13%	14	23%	24	39%	5	8%	11	18%
Fat (adjective) (46)	2+0	4%	31	67%	10	22%	1	2%	2	4%
Total (272)	25+3	10%	78	29%	120	44%	12	4%	34	13%

[1] The figures in this column indicate the number of words whose listing changes in *OALD6*. The words either make their first appearance in *OALD6* or are left out for the first time in *OALD6*. The two figures are separated by a plus sign.
[2] The figures in this column indicate the number of words which either appear and disappear, or disappear and reappear.

Table 9c. Frequency of labelling in different semantic fields

	Percentage of labelling in					
	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Prostitute	0	18	21	38	42	15
Homosexual (noun)	–	0	70	64	73	67
To have sex	0	0	29	35	22	26
Buttocks	0	22	18	13	13	12
Exclamation of anger	0	25	29	27	27	35

Percentage of labelling in					
OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Ethnic minority 11	14	42	61	63	73
Stupid person 0	0	0	31	31	20
Fat (adjective) 0	0	0	14	12	19

Table 9d. Overall frequency of labelling in the six dictionaries studied

	OALD1	*OALD2*	*OALD3*	*OALD4*	*OALD5*	*OALD6*
Words listed (272)	99	143	183	223	208	232
Words labelled +	1	9	28	58	45	55
Words labelled (+)	1	2	6	14	18	16
Percentage of labelling	2	8	19	32	30	31

Table 9e. Degree of uniformity in labelling for different semantic fields

	Words not comparable[1]	*Uniformity*		*Near-uniformity*[2]		*Non-uniformity*	
Prostitute (20)	4	10	*63%*	3	*19%*	3	*19%*
Homosexual (noun) *(14)*	3	9	*82%*	2	*18%*	0	
To have sex (42)	8	28	*82%*	3	*9%*	3	*9%*
Buttocks (19)	4	14	*93%*	1	*7%*	0	
Exclamation of anger (35)	5	26	*87%*	3	*10%*	1	*3%*
Ethnic minority (34)	2	13	*41%*	8	*25%*	11	*34%*
Stupid person (62)	8	41	*76%*	5	*9%*	8	*15%*

	Words not comparable[1]	Uniformity		Near-uniformity[2]		Non-uniformity	
Fat (adjective) *(46)*	2	35	*80%*	4	*9%*	5	*11%*
Total (272)	36	176	*75%*	29	*12%*	31	*13%*

[1] The treatment of a particular word is not comparable if the word is listed in one dictionary only.
[2] 'Near-uniformity' means that one dictionary differs in its labelling from the other dictionaries.

Table 9f. Overall frequency of the taboo sign ⚠ in the six dictionaries studied

	OALD1	OALD2	OALD3	OALD4	OALD5	OALD6
Negatively labelled words	2	11	34	72	63	71
Words labelled ⚠	0	0	8	20	16	14
Words labelled ⚠ with another label	0	5	18	9	7	24
Percentage of ⚠ sign	0	45	76	40	37	54

Table 9g. Degree of uniformity in use of the taboo sign ⚠ in different semantic fields

	Words not comparable[1]	Uniformity		Near-uniformity		Non-uniformity	
Prostitute (20)	14	5	*83%*	1	*17%*	0	
Homosexual (noun) *(14)*	6	1	*13%*	2	*25%*	5	*63%*
To have sex (42)	33	9	*100%*	0		0	
Buttocks (19)	17	1	*50%*	1	*50%*	0	
Exclamation of anger (35)	26	9	*100%*	0		0	
Ethnic minority (34)	15	12	*63%*	5	*26%*	2	*11%*

	Words not comparable[1]	Uniformity		Near-uniformity		Non-uniformity	
Stupid person (62)	50	10	*83%*	1	*8%*	1	*8%*
Fat (adjective) *(46)*	42	4	*100%*	0		0	
Total (272)	203	51	*74%*	10	*14%*	8	*12%*

[1] The treatment of a particular word is not comparable when the word is labelled negatively only in one (or none) of the dictionaries.

Appendix B: Thesauri Entries for the Semantic Field 'Homosexual'

Mawson, C. O. Sylvester (ed.) 1936 (1940). *Roget's Thesaurus of the English Language in Dictionary Form.* Revised ed., 1st ed. 1931. Garden City: Garden City Publishing Company.

> No entry for *homosexual.*

Brown, Ivor (ed.) 1962 (1966). *Roget's International Thesaurus.* 3rd, ed. 1st ed. n.d. London: Collins.

> *homosexual* 418.11
> homosexual, homosexualist, intersex, sex-intergrade, hermaphrodite, androgyne, gynandroid, bisexual, epicene, Lesbian, Sapphist, pervert, deviant, sexual pervert, sex pervert, sodomist, sodomite, pederast, homo, queer, fairy, pansy, queen, nance, Nancy, Molly, Miss Molly, betty, painted Willie, fag.

Dutch, Robert A. (ed.) 1962 (1977). *Roget's Thesaurus of English Words and Phrases.* New revised and modernised ed., 1st ed. 1952. London: Longman.

> *homosexual* 84 *n.* (nonconformist)
> homosexual, lesbian, pansy, fairy, queer, pervert; sadist, masochist.

Kirkpatrick, Betty (ed.) 1987. *Roget's Thesaurus of English Words and Phrases.* New ed., 1st ed. 1852. London: Longman.

> *homosexual* 84 *n.* (nonconformist)
> invert, homosexual, lesbian, les, lez, gay, pansy, fairy, nancy, poof, poofter, homo, fruit, queen, queer, dyke, dike, transvestite, drag artist, gender-bender; …

Urdang, Laurence (ed.) 1991. *The Oxford Thesaurus: An A-Z Dictionary of Synonyms.* Oxford: Clarendon.

> *homosexual n.*
> gay, homophile, lesbian, tribade, sapphist, pervert, invert, queer, fairy, pansy, nancy, nance, queen, drag queen, homo, butch, (bull) dyke, poof, poofter, ginger (beer), fruit, auntie, fag, faggot.

Chapman, Robert L. (ed.) 1992 (1996). *Collins Roget's International Thesaurus.* 5th ed. 1st. ed. 1852. Glasgow: HarperCollins.

> *homosexual n.* gay person 75.14 (sex)
> gay person, homosexualist, homophile, invert, catamite, mignon, Ganymede, punk, gunsel, bisexual, bi-guy, lesbian, sapphist, tribade, fricatrice.
>
> *n.* informal terms for male and female homosexuals 75.15
> homo, queer, faggot, fag, fruit, flit, fairy, pansy, nance, auntie, queen, drag queen, closet queen, poof, poofter, poove, dyke, lesbo, lez.

Ferguson, Rosalind, Martin Manser and David Pickering (eds.) 2000. *The New Penguin Thesaurus.* London: Penguin Books.

> *homosexual*, gay, lesbian, queer, poof, poofter, fairy, queen, homo, dyke, faggot, fag, bent, pansy, nancy, woofter.
>
> homosexual, lesbian, pansy, fairy, queer, pervert; sadist, masochist.

Appendix C: Inventory of Usage Labels as Listed in the User Guides

OALD1 and *OALD2* do not provide a list of usage labels.

OALD3:
Specialist English registers:

accounts	*electricity (electr)*	*phonetics (phon)*
aerospace	*farming*	*photography (photo)*
algebra (alg)	*finance (fin)*	*physics (phys)*
anatomy (anat)	*football*	*physiology (physiol)*
architecture (archit)	*gambling*	*politics (pol)*
arithmetic (arith)	*geology (geol)*	*psychology (psych)*
art	*grammar (gram)*	*racing*
astronomy (astron)	*history (hist)*	*radio telegraphy (radio)*
ballet	*journalism*	*rugby*
biblical	*legal*	*science*
biology (biol)	*linguistics (ling)*	*sport*
book-keeping	*mathematics (maths)*	*tennis*
botany (bot)	*mechanics (mech)*	*theatre*
business	*medical (med)*	*trigonometry (trig)*
chemistry (chem.)	*meteorology (met)*	*zoology (zool)*
cinema	*military (mil)*	
commerce (comm)	*music*	
computers (comp)	*mythology (myth)*	
cricket	*nautical (naut)*	
ecclesiastical (eccles)	*pathology (path)*	
engineering (eng)	*philosophy (phil)*	

Stylistic values:

archaic	*formal*	*poetic (poet)*
colloquial (colloq)	*humorous (hum)*	*proverb (prov)*
dated	*ironical (ironic)*	*rare*
derogatory (derog)	*jocular (joc)*	*rhetorical (rhet)*
dialect (dial)	*laudatory (laud)*	*slang (sl)*
emotive (emot)	*literary (liter)*	*vulgar (vulg)*
emphatic (emph)	*literally (lit)*	⚠ (taboo)
euphemistic (euphem)	*modern use (mod use)*	
facetious (facet)	*old use*	
figurative (fig)	*pejorative (pej)*	

OALD4:
Currency:
dated, *arch*
Region:
Brit, *US*, *Scot*, *dialect*, *S African*, *Austral*, *NZ* and other labels spelled out in full
Register:
⚠ (taboo), *sl*, *infml*, *fml*, *rhet*
Evaluation:
derog, *approv*, *offensive*, *euph*, *ironic*, *fig*, *joc*, *sexist*
Technical field: "mostly self-explanatory", for example:
architecture, *law*, *art*, *cinema or TV*, *computing*, *grammar*
Sayings and catchphrases: *saying*, *catchphrase*
Proprietary names: *propr*

OALD5:
Labels indicating words that express a particular attitude or are appropriate in a particular context:

approv	*infml*	*sexist*
derog	*ironic*	*sl*
euph	*joc*	⚠ (taboo)
fig	*offensive*	
fml	*rhet*	

Labels indicating other restrictions on the use of words:

arch	*dialect*	*techn*
Brit	*propr*	*US*
catchphrase	*saying*	
dated	*Scot*	

Labels indicating a restriction to specialized use in a specific field, for example:

anatomy
computing
grammar
law

OALD6:
Labels indicating words that express a particular attitude or are appropriate in a particular situation:

approving	*ironic*	*technical*
disapproving	*literary*	*written*
figurative	*offensive*	⚠ (taboo)
formal	*rare*	
humorous	*slang*	
informal	*spoken*	

Labels indicating other restrictions on the use of words:

AmE	*saying*
BrE	™ (trademark)
dialect	
old-fashioned	
old use	