

INTRODUCTION

THIS new edition of the *Oxford English Dictionary* contains the whole text, unaltered in all essentials, of the twelve-volume first edition, which appeared in 1933 as a reprint of the ten-volume *New English Dictionary on Historical Principles*, itself originally published in parts between 1884 and 1928. It also contains the complete text of the four-volume *Supplement to the Oxford English Dictionary*, published between 1972 and 1986; this superseded the previous *Supplement*, which was issued in 1933 as a companion to the main work.

The main purpose of this second edition is to present a version of the Dictionary in which these two parts, the twelve volumes and the four volumes, are amalgamated into a continuous, seamless text. Accordingly, every article from the *Supplement* has been either added in its proper alphabetical position (if a wholly new entry) or merged with its corresponding *OED* entry (following directions which the *Supplement* gives but the present edition omits, as now redundant). Instead of the sixteen volumes of large but unequal size in which the previous works were issued, it has been published in twenty slimmer and evenly sized volumes, which it is hoped will prove more convenient to use.

Although the *raison d'être* of this new edition is the integration of the two texts, and adherence to the instructions of the *Supplement*, whether explicitly stated or contextually implied, is the guiding principle of the work, the material brought over from the *Supplement* is by no means the only feature that differentiates the second edition from the first. New vocabulary has been added, certain important general revisions, and numerous local corrections, have been made, and the whole text has been given an entirely new typographical format. It is estimated that these changes, fuller details of which are given below, have affected (in different ways) the majority of the 290,500 entries contained in this edition, including virtually all articles on the commonest words. Together they have made this edition significantly richer in information, and more modern in aspect, than its distinguished predecessors.

Whereas the *Supplement* can be regarded for practical purposes as up to date, it is a matter of common knowledge that many elements of the original *OED* require revision. That is the very purpose for which the New OED Project, of which the present work is the first printed product, was initiated. Several of these requirements have been addressed in this edition. But the full revision and updating of the Dictionary (an outline agenda for which is given in the *History* section below) must be regarded as a long-term goal, demanding considerable resources, and therefore to be approached in stages. This new edition represents the first, and almost certainly the most arduous, step towards that goal.

The fundamental difference between this edition and its predecessors is, by its nature, quite invisible to the user of the Dictionary. Before the compilation of this edition, the *OED* and *Supplement*, which had only ever been typeset in hot metal, were computerized. The machine-readable version of the Dictionary resulting from computerization is now the master copy from which the present printed edition has been made. It is the version to which further modifications to the Dictionary will be applied, and from which new editions and offspring works of reference will be generated in the future.

But the machine-readable text is not distinct from the printed one merely by virtue of existing in a different physical medium, electronic instead of paper, or tape rather than type; or even by its priority in the production process. Much more important, it carries a whole new world of information. This has nothing to do with supplementary text. In addition to the conventional natural-language text taken over from the printed *OED* and *Supplement*, there is another layer of information: the

'mark-up language' or tagging system. The different categories of information into which the text can be partitioned, such as headword, pronunciation, variant form, quotation, or date of quotation, are each identified accurately and unambiguously by computer tags. These tags consist of short sequences of letters, mainly mnemonic in form (e.g. 'quot' for 'quotation'), set off by delimiting characters from the conventional text which they mark. They do not appear on the printed page, but are there translated into various features of layout, typography, and punctuation. Unlike the tags, most of the latter, viewed logically, are to some degree ambiguous and redundant, though familiar and convenient to the user of the Dictionary. Thus, the tags that signify 'headword' and 'date of quotation' accompany these elements, but no instruction to print them in bold type appears in the machine-readable text; they could as easily be printed in capitals or sanserif if desired. The typographical realization of an element is relative, being determined by the requirements of the particular published form; accordingly, information of this kind has been largely, though not absolutely, purged from the electronic master text.

Only the presence of this structural information within the text has made it possible to integrate the *OED* and *Supplement*, and to perform the other systematic changes listed below, with so small a staff and in so short a time, and only because of it is the further revision and updating of the Dictionary feasible at all. It is also the prerequisite for the conversion of the *OED* into a publicly available electronic database. The project team believes the addition of this information to have been, without doubt, their chief contribution to the future of the *Oxford English Dictionary*.

SPECIAL FEATURES OF THE SECOND EDITION

The distinctive features of this edition may be described under four headings: supplementary text, general revisions, local corrections, and typographical format. These will be explained in turn.

A. SUPPLEMENTARY TEXT

1. The 69,372 entries of the four-volume *Supplement* have been amalgamated with the 252,259 entries of the *OED*, first edition. 41,752 of these entries are new and independent; the remaining 27,620 have been integrated with the corresponding *OED* entries. The principles that guided this process of integration are explained below.
2. Entirely new articles dealing with an additional 5,000 words, combinations, and senses, have been included and integrated; these are located chiefly in the first third of the alphabet, where the work done for the *Supplement* is now twenty years or more old. The policy and history of this part of the project are set out below.
3. The 260 addenda and 83 spurious entries appended to Volume XII of the first edition have here been merged with the main text.
4. 560 corrections, being chiefly earlier illustrative examples, which were prepared for Volumes I and II of the *Supplement* but not inserted there, have been included here.
5. The process of integration has from time to time required that a lexical item, treated as a subordinate part of an entry in the *OED* or *Supplement*, should be elevated to the status of a main entry, and this has naturally entailed the writing of new text.

B. GENERAL REVISIONS

1. *IPA and stress-marked headwords*. The system devised by Sir James Murray for representing pronunciation, used in both the first edition of the *OED* and the *Supplement*, has now been replaced

throughout the text by the International Phonetic Alphabet. Many headwords and lexical items in the two parent works had their stress-pattern marked by symbols placed within them, instead of being followed by a phonetic transcription; these marks, which are placed *after* the stressed vowel, have been replaced by IPA stress-marks, which are placed *before* the stressed syllable. The principles of transcription and translation followed here are described below, pp. xxxiii–xxxiv.

2. *Foreign script.* In its etymological material, the first edition regularly cited foreign words in non-roman scripts; besides Greek, cited forms in Arabic script, the Cyrillic alphabet, Devanagari, the square Hebrew alphabet, and the Syriac script are quite usual. These were normally, but not universally, accompanied by transliterations. Except when citing Greek, the *Supplement* abandoned this practice, giving only transliterations. It was decided to follow the latter's practice in the present edition, considering that the dropping of the scripts would be more straightforward than the furnishing of accurate new non-roman forms, and that the first edition itself frequently neglected to supply the non-roman forms. Transcriptions have been supplied wherever they were missing in the first edition.

In quotations the presence of foreign script is, of course, an intrinsic feature; it has been preserved as far as possible, subject to the constraints upon 'artwork' and special characters in general (see below).

3. *Illustration and special characters.* The parent texts resort from time to time to the inclusion of what amount to pictorial illustrations, mainly diagrammatic or typographic in style. In the first edition a number of names for typefaces are typographically illustrated, and a few other concepts are conveyed diagrammatically. These have been omitted. Other more modest forms of illustration, which involve the use of individual special characters such as occur or might reasonably be expected to occur in the Dictionary, have been retained.

Between them, the two parent texts make use of approximately 660 characters apart from the ninety or so available on the typical keyboard. Virtually all of these have been retained, and some previously wrong have been corrected.

4. *Ordering of entries.* The alphabetical arrangement of entries in the *OED* and *Supplement* is to some extent affected by the presence of special characters, accents, punctuation, and capitalization within the headword. The principles which prevail, but are not universally followed, in the parent texts have been standardized throughout the present edition. As a result, certain details in the identification of some entries differ from their counterparts in the parent texts, and a few of these have consequently been removed some distance from their former position.

5. *Ordering of senses.* The sense-divisions of most entries in the first edition and its *Supplement* follow a very clear system of structural organization, as described below, p. xxxiii. The system has been extended to the few scattered entries which were (usually for no special reason) irregular in structure.

Entries in which a series of senses skips or duplicates a number, owing to simple editorial or typographical oversight, have been corrected.

6. *Cross-references.* Cross-references whose targets were changed as a result of the integration of *OED* and *Supplement* entries have been emended as far as possible. These changes reflect the changes to the identifying structure of an entry, listed below, p. xviii.

Many of the 580,000 cross-references in the Dictionary are imprecise, citing headwords without parts of speech and homonym numbers, for example. It was impossible for the automatic cross-referencing system to determine which of two or more possible targets was the one proper to an

ambiguous cross-reference of this sort, and so, on the whole, these have not been made more precise; in many cases, the intended target is obvious to the reader, and amplification would merely be fussy. There were also a fair number of cross-references which, as printed, did not match any existing headword; this was nearly always because of a slight difference in spelling. Most of these have been emended in the present edition.

C. LOCAL CORRECTIONS

1. *The spelling of vocabulary items.* Certain conventions of spelling, as also of capitalization, hyphenation, and punctuation, have changed since the publication of the first edition; indeed the occurrence of some such changes is evidenced within the Dictionary itself. Harmonization of the whole text with currently acceptable style would have been impossible within the limits of this new edition. The *Supplement*, however, indicated many changes to the spelling of headwords, which have, of course, been effected; and an attempt has been made to carry such changes through into derivatives and combinations of the main words and into contiguous definitions. Other such updatings, overlooked by the *Supplement*, are carried out wherever possible.

2. *The main text of Dictionary entries.* Innumerable small misprints and slips have naturally been encountered, during editing, in the definitions, etymologies, and notes which form the core of the Dictionary text. These have been corrected.

3. *Quotations.* The text of quotations has been carefully protected from corruption. The working assumption was that it always correctly reproduces the original source, however strangely it may read. Nevertheless, an appreciable number of quotations came under suspicion of inaccuracy, or could be clearly seen to have suffered mutilation at the hands of compositors, and were checked and corrected from the sources.

It was a basic, and not unreasonable, requirement of our automatic processing that quotations (with certain regular exceptions, such as those from *Beowulf*) must begin with a date. Dates (sometimes only approximate) were supplied by means of bibliographical investigation to the small number of quotations that were found to lack them.

D. TYPOGRAPHICAL CHANGES

1. *Entry spacing.* In the first edition, no spacing separates entries one from another. This edition follows the *Supplement* in placing space between entries. Series of entries for variant and obsolete forms are treated in the same way, not run on as they often are in the first edition.

2. *Distinction between main and subordinate entries.* The typographical distinction in the first edition between main words and subordinate words, by which the latter were printed in a lighter bold type, has been given up in this edition, as it had already in the *Supplement*. Such a distinction is difficult to draw absolutely and is, in any case, of doubtful utility.

3. *Distinction between headword and other bold elements.* Besides the distinction between two kinds of headword just described, the first edition used other varieties of bold type to identify derivatives, combinations, and variant forms of the headword, when cited within the same entry. Derivatives were usually printed in dark bold similar to, but smaller than, that of the headword, while combinations and variant forms were printed in a lighter bold. The *Supplement* used only a single typeface to distinguish all three from the headword. In this edition derivatives and combinations are printed in a dark bold, smaller than the headword, and variant spellings are printed in a light bold. It seemed

logical to symbolize in the same way what are, in effect, subordinated headwords, but to differentiate them from the variant spellings of the main headword.

4. *Italicized vocabulary items.* The text of a Dictionary definition contains numerous elements printed in italics, which fall into several different categories of information: chiefly usage label, cross-reference, cited linguistic form, and lexical item (such as phrase or minor combination). This last element is of particular importance since, like the headword, derivative, or bold combination, it constitutes one of the keys by which the reader finds the information which he or she is seeking. Since this kind of element is specially marked by tags in the electronic version of the text, it seemed helpful to print it in this edition in a special bold italic typeface, clearly setting it off from all other italicized text.

5. *Capitalization of headwords.* In the first edition of the *OED*, every main headword was given a capital initial, regardless of whether the word was normally so written. Most derivatives, and many combinations, were also capitalized. The *Supplement*, in accord with modern lexicographical practice, abandoned this convention, giving a capital only where that is the normal spelling. This edition follows the *Supplement's* practice.

For many words capitalization varies, either at different dates or in different senses. Because its convention disguised the problem, the *OED* often did not indicate the prevailing or preferred style. Where the intentions of the first edition were not deducible, as often with rare and obsolete words, decisions about capitalization were made on the basis of the printed quotations or analogy with similar and related words, or both.

6. *Abbreviations in initial letter entries.* Only a small number of abbreviations (i.e. initialisms) were listed under the entries for initial letters in the first edition. In line with recent linguistic developments, these lists were greatly augmented by the *Supplement*. But though these abbreviations have definitions, they are not picked out typographically in the parent texts. In this edition they are printed in bold type for easy identification.

7. *Asterisks in quotation paragraphs and cross-references.* In the Dictionary, quotations illustrating a series of combinations can be listed either in one chronological sequence, or (as is usual when the combinations are defined) in a series of chronological sequences, each illustrating one combination and all arranged in the alphabetical order of the combinations. In the first edition, an asterisk was placed in the first quotation of each sequence, marking the combination being illustrated. In the *Supplement*, this convention was not followed, because asterisks were used to mark two other features: sense numbers that were to be intercalated into the *OED* sequence, and cross-references to entries in the *Supplement*. As a result of integration, the latter conventions have disappeared; but also, many quotation series without asterisks from the *Supplement* have been merged with series with asterisks from the *OED*. In these cases, the asterisking convention has been carried through the whole merged quotation paragraph. It has not, however, been introduced into every paragraph of this kind originating in the *Supplement* alone.

8. *Hyphenation.* Unlike its parent texts, this edition has been printed without regular line-end hyphenation. Most of the hyphens printed are true ('hard') hyphens. This has the advantage that no extraneous hyphens are introduced into lexical items, variant forms, or other linguistic forms cited in the text. It also means that virtually no merely line-end ('soft') hyphens have been introduced into the text of quotations. Though this results in a less even layout of text on the page than in the parent texts, it is felt that the advantages outweigh this drawback.

When the text of quotations from the Dictionary was keyboarded, hyphens occurring at line

endings had to be either dropped or retained (as ‘hard’ hyphens). Without consulting the original works from which the quotations were drawn, it was sometimes impossible to decide which would be correct, even after considering the date of the quotation, the evidence of the other quotations from the same work, and so on. In order to avoid a misleading decision, a special symbol (-) has been used to replace the hyphen of the parent text. This symbol indicates nothing more than the ambiguity of the hyphen in the parent text. It is also occasionally used to split a bold or italic combination, a derivative, or a word employed in a definition, for the sake of a line-break: in such cases, it is to be understood to indicate that the word is *not* normally written with a hyphen.

PRINCIPLES OF INTEGRATION

The integration of the text of the *Supplement* into that of the first edition of the *OED* (referred to below as ‘the *OED*’ for brevity’s sake) was carried out in two stages. The major processes were performed automatically by computer programs specifically developed for the purpose. Printouts of the resulting merged entries were then edited by lexicographers, and the emendations marked on them were entered by keyboarders into the computer.

The guiding principle of both stages of the integration process was that the intentions of the *Supplement* should be faithfully followed unless there were very good reasons for departing from them.

INDEPENDENT ENTRIES

About 42,000 main entries in the *Supplement* were new and independent. They were put into their appropriate place in the alphabetical sequence.

A new entry having a headword with the same spelling and part of speech as one or more in the *OED* already had a different homonym number: it was placed after its *OED* homonym(s). If there was only one homonym in the *OED*, this was given the homonym number 1. A new entry having a headword with the same spelling but a different part of speech was placed in an appropriate position among the pre-existing entries, usually towards the end of the sequence (see the principles of entry ordering, p. xxxii below); where only a noun (‘substantive’) or nouns occurred in the *OED*, they had no explicit part of speech: the part of speech ‘*sb.*’ was supplied.

Many combinations and derivatives treated, in the first edition, within the entries for the main words on which they are formed were elevated to the status of main words in the *Supplement*. The usual reason for this was a significant increase in the complexity of the senses of the word or its general acceptance as more than a casual or obvious compound of the root word. A certain number of forms treated as graphic or spoken variants of a main word in the first edition were similarly registered as separate, independent words in the *Supplement*. When such upgrading occurred, the material relating to the word in the first edition was not reprinted in full in the new entry; instead, cross-references were used to direct the reader to the definition, etymology, or quotation in question, implying that it was now part of the new entry. Accordingly, in the integrated text this original material has been moved (with appropriate editing) to the new position. Much effort has been devoted to detecting overlapping entries of this sort, many of which are extremely well disguised, especially when linked by only a single, inconspicuous quotation.

A certain number of *Supplement* entries directed the reader to delete the existing Dictionary entry and substitute a new one, the text of which then follows. Not infrequently, these entries also borrowed snatches of text from the former entry, especially the quotations given there. Before the original entries were discarded, the borrowed text was transferred to the new entries.

There were also entries in the *Supplement* which added further information to a different main word in the first edition, but nevertheless did not require a new entry to be created. There were two principal kinds. First, where the entry for a main word in the first edition had a derivative listed under it and the only addition to that entry in the *Supplement* related to the derivative, the latter was given, as it were, the temporary status of a main word solely for the purposes of supplementation. In such cases, the new material has been accommodated within the framework of the original main entry and the supplementary one has been discarded. Secondly, a variant form of a main word could be treated in a separate entry with illustrative quotations, because the variant in question either had not been listed and illustrated under the principal form, or had recently become recognized as the predominant spelling. In both cases the intention was that the two entries should be merged, in the latter with a change of spelling in the headword. Again, detecting such pairs of entries (and some forming larger groups) has been very difficult, especially as the spelling variation can bridge widely separated parts of the Dictionary (C and K, or E and Ē, for example), a factor which has occasionally frustrated the fulfilment of the *Supplement*'s intentions.

MATCHING ENTRIES

The remaining 28,000 entries in the *Supplement* had counterparts in the *OED* with which they were integrated. Four basic procedures were followed, in line with the main instructions to the user included in each entry: *addition*, *deletion*, *transference*, and *substitution*. Naturally not every one of the thousands of occasions when any of these changes was made is explicitly indicated in the text, but the intentions are nearly always quite clear.

In the headword section of the entry (the *identification*), the usual supplementary material consisted of a pronunciation or one or more variant forms or inflexions. Deletions were mainly of status labels (especially 'obsolete'). Transference of labels into a sense section was common. Substitution of a different spelling of the headword or a modified label occurred from time to time.

Etymologies were occasionally added or modified by deletion or substitution.

In the *signification*, additions made up the bulk of the changes occasioned by the *Supplement*. Where a complete new sense section was added, its numbering normally indicated the correct position. If the first edition had only one sense section at that level in the hierarchy, the absent sense number or letter (the first in the series: A, 1, a, or (a)) was supplied. *Supplement* sense sections introduced by starred numbers were added at the appropriate point in the sequence, and the ensuing sense divisions were renumbered to accommodate them. When a section division (headed by a capital letter), covering the use of a word in a new grammatical category, was added at the end of an entry, the identification was altered to include the new part of speech. Sections containing derivatives were added at the end of the entry.

Supplementary material was frequently added to definitions; this was commonly made plain by the introductory word 'also', which sometimes became superfluous and was edited out. Occasions when the new material could merely be added, without modification, at the end of the existing definition were relatively rare. Often the new material had to be inserted within the old, or the whole rewritten into one definition. The domain of a label had sometimes to be restricted to cover only the intended portion of the definition. So multifarious were the modifications that the results can scarcely be distinguished from the deletions, transferences, and substitutions indicated from time to time by the *Supplement*.

The parts of the Dictionary which the *Supplement* especially augmented are the sections in which defined and undefined combinations are listed. Considerable labour was required to merge the corresponding lists (usually in alphabetical order). The defined combinations, resembling as they do complete entries in miniature, were susceptible to supplementation, deletion, modification, and

substitution in any of their constituent parts. Their alphabetical sorting and rough merging was achieved automatically, but much work remained for editorial attention.

The addition of *quotations* was very frequent: when an entirely new sense-division was added, with its own group of quotations, when a sense section was supplemented, and as the sole modification to a sense section. In most cases this was a straightforward operation. Paragraphs of quotations that illustrate combinations, however, usually consist of a series, arranged alphabetically to correspond with the combinations, of short chronological sequences of quotations; the accurate fusing of parallel series of these was a complicated operation, requiring careful editing, both on paper and at the screen.

CROSS-REFERENCES

Integration of the *Supplement* with the main Dictionary text involved an enormous number of changes to the 'addresses' at which sections of text were located: namely, to the form of the headword, the part of speech associated with it, the homonym number distinguishing otherwise identical main words, or the number or letter identifying the sense-division under which a definition, combination, or quotation could be found. As a result, there was a danger that many cross-references would become invalid, since the elements to which they pointed would now be differently identified. This was countered by a twofold strategy. A computer program caused every change to the 'address' of an element brought about by the process of automatic integration to be applied to the cross-references which cited that address. Editorial staff subsequently ensured that changes made by them were logged and then applied to the corresponding cross-references. Inevitably, a few cross-references escaped both systems, including those in early volumes which were already passed for press before editorial changes that would affect them had been made in later volumes. There are around 600,000 cross-references in the *OED*, of which well over 20,000 have been adjusted in response to integration, and many more have been corrected or rendered more precise.

THE TRANSLATION OF THE PHONETIC SYSTEM

The system of phonetic transcription devised by Sir James Murray for use in the first edition and followed, for the sake of consistency, in the *Supplement* is a subtle and flexible means of recording English pronunciation. But many of the effects for which Murray strove in the design of his system were realized, not long afterwards, in the International Phonetic Alphabet (IPA). It seems very possible that, if the IPA had already achieved full development and widespread acceptance at the time that Murray was beginning work on the Dictionary, he would have adopted it instead of a system of his own. It was therefore logical to consider replacing Murray's system with IPA throughout the Dictionary. IPA has the advantage that it is very widely accepted and understood, and can be used to represent the sounds as well of regional and dialect English and foreign languages as of standard English. Indeed its introduction was regarded by many whom the project team consulted as among the highest priorities. It was decided that the change should be made for the present edition, rather than left until a future revision phase. But for reasons of historical interest, the Murray transcriptions have been retained in the electronic version alongside the IPA ones, the latter only having been printed.

A notable feature of the Dictionary is that obsolete main words, derivatives, and certain variant forms and combinations are not given phonetic transcriptions but have their stress-pattern indicated: the stress-dots (which are placed *after* the accented *vowel*, just as they are within the transcriptions) are printed within the body of the word or form. Naturally consistency required that the stress-dots within these forms should also be altered to IPA stress-marks (which are placed *before* the beginning of the accented *syllable*).

The short time available meant that there had to be rather strict limitations on the extent of the changes made. For native English words the variety of pronunciation represented is, broadly speaking, educated standard southern British, or 'Received Pronunciation' (RP). There could be no question, at this stage, of systematically registering non-RP (i.e. other British and non-British) pronunciations; although of course those already included in the Dictionary have remained and have been augmented by a few analogous cases. Again, pronunciation variants within RP, apart from those already registered in the Dictionary, could not be researched and documented, except as the evident need to include them arose in the course of the other editing. The adding of up-to-date pronunciations, except when prescribed by the *Supplement*, had to be kept to a minimum, for the same reasons. Essentially, a straightforward literal translation from the Murray system to IPA has been attempted, accompanied by correction of the errors inevitably arising from that process.

The method by which the translation of the Murray system to IPA was performed can be briefly summarized. As much of the translation as possible was carried out automatically by computer. The computer had to operate on two kinds of material: sequences of phonetic characters, and normal English words containing dots indicating the position of the stress accent. First, the program identified the strings of characters which it was required to process. It used the mark-up tags to do this. Then the actual translation was run. The program referred to three tables, one giving IPA equivalents for Murray's phonetic symbols (or groups of them); one giving rules that showed whether each IPA symbol is a vowel, a consonant, or a consonant cluster; and one giving similar rules for ordinary English spelling. The second and third table enabled the computer to deduce the correct position for an IPA stress-mark from the existing stress-dot in the text. Certain characters and groups were known in advance to have more than one possible IPA equivalent. These were printed out on a special report, as were all pronunciations in which the conversion failed; and from these, corrections were made at the keyboard. Altogether, 137,152 phonetic transcriptions and 137,274 stress-marked words were automatically translated.

A considerable number of phonetic transcriptions are found in other parts of an entry than the pronunciation key. These were not originally identified by the computer, and so a separate program was run in order to register them on a list for editorial transcription.

The conversion of Murray's phonetic system to IPA was not entirely straightforward. The peculiar nature of the system, the way in which it is applied both in general and in particular cases, and the historical variety of English speech which it was employed to record, all presented obstacles to the smooth application of the scheme of translation.

The peculiarities of Murray's phonetic alphabet can be ascribed to the time and circumstances of its devising.¹ It must here suffice to enumerate a few of its characteristics. It is used indiscriminately both for *phonetic* and *phonemic* transcription: that is, it is a set of symbols employed to represent both the members of the particular set of sounds that constitute the phonology of English, and the much larger set of infinitesimally differing sounds from all other phonologies to which reference is made in the Dictionary. As a means of representing standard English pronunciation, Murray's system is sensitive and generally lucid. It is less well-adapted for the transcription of dialectal and foreign words. Shift had often to be made with the limited range of typographical shapes that Murray had devised at the outset in order to transcribe unexpected foreign sounds. Symbols not listed in the original pronunciation key appeared during the progress of the Dictionary. Uncommon foreign sounds were handled differently on widely separated occasions.

With regard to consonants, there are few drawbacks to Murray's system, apart from the dearth of symbols for foreign sounds. Indeed, all but one of the primary consonant symbols correspond exactly to the set employed for English in the IPA. A fundamental feature of the vowel system that does not

¹ For a full discussion, the reader is referred to M. K. C. MacMahon, 'James Murray and the Phonetic Notation in the *New English Dictionary*', *Transactions of the Philological Society*, 1985, pp. 72-112.

translate easily into IPA is its analysis into ‘ordinary’, ‘long’, and ‘obscure’ vowels. Broadly, the same list of symbols appears under each heading, but (respectively) without a diacritical accent, accented with a macron (long mark), and accented with a breve (short mark). The implication is that the same underlying vowel manifests itself in three guises; and, further, that these guises are determined at least partly by its relationship to the stress accent of the word. The theory underlying the obscure vowels is that, *if* they were accorded stress, as they sometimes are in song or very careful enunciation, they would resemble their ordinary equivalents. This principle, though something like it is encountered in phonological theories, is not usually embodied in the standard IPA transcriptions used in dictionaries. If, as has here been done, it is set aside, then most of the obscure vowels are equivalent to the IPA ə (‘schwa’; the sound of *a* in *sofa*, *particular*); while a few are translated into ɪ (the sound of *e* in *hatchet*). The main peculiarity of the long vowel series is that diphthongs are included in it; but this is an oddity of phonetic theory—the analysis of English diphthongs as long vowels followed by ‘glides’—rather than an obstacle to translation.

The major problem arises in the ‘ordinary’ vowel series, which contains both vowels universally recognized as lacking length—the *i* of *pin*, for example—and vowels now normally transcribed as long vowels. These, in fact, are the members of the long series in another guise: a guise apparently taken on in a syllable under low stress. So, for example, the vowel of the second syllable of *Matthew* is an ordinary vowel, while that of *few* is long. Most classifications would treat them as identical. This would not pose a major problem of translation, if it were not that the same symbols employed for the second vowel of *Matthew* and its parallels throughout the vowel system (*hero*, *Psyche*, etc.) are also used to represent the principal vowels of the European languages—roughly speaking, the so-called ‘cardinal vowels’. The two sets of vowels are not phonetically identical, or even close, and were not in Murray’s time. Such symbols as these (along with some others) were straightforwardly ambiguous, and could not be translated correctly in every case by the computer.

Because Murray’s system uses many of the same symbols for English and foreign vowel sounds, it was necessary to depart from certain conventions that are common in the transcription adopted by English dictionaries. A notable example is the use of (ɛ) for the vowel of *bed* in place of the more usual (e), because the English vowel is, for most English ears, closer to the cardinal open vowel of French *faire* or German *Bär* than to the cardinal narrow vowel of French *bébé* or German *Schnee*.

The variety of English pronunciation recorded by the Murray transcription is extremely ‘precise’, conservative, and (in present-day terms) old-fashioned. Most of its peculiar characteristics are systematic (they permeate the phonology) rather than occasional (features of the pronunciation of particular words, or groups of words). They are systematic enough for a phonetician to predict the way in which individual words will be transcribed, but not enough to make it easy for a computer to efface them all automatically. This is one reason why they have largely been retained in this edition. A second reason is that they constitute a useful record of one variety of English pronunciation in a particular period; and a third is that, for the general user, most of them are merely small nuances for which one can make allowance.

The chief phonological features that set the pronunciation represented in the Dictionary apart from most present-day phonetic descriptions of English are the following:

1. The stressed close vowels (i:, u:) are maintained in words like *idea*, *realize*, *museum*, *skua*, in contrast with the diphthongs (ɪə, ʊə) of *dear*, *rear*, *secure*.
2. The close vowels (i:, u:) are maintained in unstressed syllables in words like *delineate*, *creation*, *perpetual*, *graduate*, in contrast with the open vowels (ɪ, ʊ) of *genius*, *demonic*, *circulate*.

It is in such syllables that the Murray transcription uses ‘ordinary’ vowels, not long ones. It is quite likely, therefore, that Murray was aware of no distinction in duration between these vowels,

but it is also clear from the consistency with which the distinction is observed that he was aware of, or believed in, a distinction of vowel quality (specifically, of tension). The conventional IPA transcription, however, observes only a phonemic distinction between open vowels, which are treated as 'short', and narrow vowels, which are accorded symbols of length as well. The IPA transcriptions therefore convey the misleading implication that such vowels have, or had, greater duration.

3. Diphthongs (e.g., əʊ) are maintained in unstressed syllables in words like *homographic*, *protocol*, in contrast with the obscure vowel 'schwa' in words like *homonym*, *melody*.
4. Unstressed medial obscure vowels are not represented as having been elided, e.g. in *veterinary*.
5. A distinction is maintained between syllabic consonants and a sequence of obscure vowel and consonant, so that words like *principle* and *principal* are distinguished.
6. A distinction is drawn between a diphthong (ɔə) in words like *glory*, *boarder*, *mourning*, and a long vowel (ɔː) in words like *saurian*, *border*, and *morning*. This is not generally maintained in present-day IPA.
7. The nineteenth-century lengthening of the rounded low back vowel of words like *soft*, *cloth*, and *cross* is recorded (in this edition, as a separate variant; in the first edition, by a special symbol indicating the possibility of either pronunciation).

NEW VOCABULARY

HISTORICAL BACKGROUND

The decision to include in this edition some five thousand new items of vocabulary was taken in the light of several considerations. First, the pace of lexical innovation is rapid, and every dictionary must be continually enlarged or revised to accommodate the many new words and senses which become part of the language each year. Secondly, the compilation of new entries is a discrete editorial activity on which work could start straight away, independent of the revision which is envisaged as the business of subsequent phases of the project. Thirdly, as work on the *Supplement* came to an end, staff who were skilled in the compilation of *OED* entries, and a research base, in the form of a large quotation file, library researchers, consultants, etc., became available. Fourthly, other Oxford dictionaries, notably the *Shorter OED*, the *Concise Oxford Dictionary*, and the *Oxford Advanced Learner's Dictionary*, then in the process of revision, needed a selection of fully-researched new entries; this could be produced most efficiently by a central unit supplying material both to the New *OED* project and to other dictionaries.

Editorial work on new vocabulary began in 1983. It was recognized at the outset that the bulk of the material would lie in the early section of the alphabet. The lexical changes of the period (between ten and twenty years) since this part of the *Supplement* was prepared were yet to be recorded; moreover, the *Supplement's* coverage of the vocabulary already in use at that period was limited both by the relatively modest scale on which the work was initially planned and by the extent of the collections, which were inevitably small in the early stages. In this it resembled, though for different reasons, the earlier *Supplement* of 1933, of which R. W. Chapman wrote: 'We projected accordingly . . . a wedge-shaped Supplement, designed to bring the work to an even date; it is, so to speak, all ABC and no XYZ.' But this asymmetry was a condition of the material being treated; there was no constraint on the editorial operation that confined it to one part of the alphabet. On the contrary, the storage of the material on the *New OED* database meant that very recent coinages from any point in the alphabet could as readily be accommodated.

SELECTION AND EDITING OF NEW MATERIAL

The large quotation file maintained by the Oxford English Dictionaries Department was the basis for selecting items for inclusion in this edition, as it had been both for the first edition and for the *Supplement*. It offered, as before, the most comprehensive and objective basis for the selection of vocabulary items and the provision of supporting evidence. The file was growing at the rate of some 120,000 quotations per year, collected predominantly by means of a reading programme covering books, magazines, etc., representing the main varieties of English used throughout the world, but enhanced by scholarly contributions of specific new information, earlier illustrative quotations, and other lexicographical material.

But new resources had become available to lexicographers during the previous decade. Perhaps the most important were the large computer databases containing research abstracts, newspaper and periodical texts, and legal reports. At first, computer-readable texts were used mainly to generate concordances. By the 1980s it was also possible, by rapidly searching the extensive text held on these databases, to pin down specific uses of individual words or phrases. These databases proved invaluable in tracing early uses of terms, in providing examples to complete the lexicographical record, and in supplying etymological information.

The databases of the 1980s rarely helped in the actual selection of new words for inclusion in the Dictionary. Inclusion was normally suggested by other means (such as evidence in the quotation file); the databases were used subsequently to provide additional information and to confirm the currency of the word or sense. It should also be stressed that these resources had to be used with some caution, as they might contain only recent text, or examples of English from only one area (notably North America), and they lacked the ability to identify the meaning in which a word was used in a given context or to distinguish it from the other meanings of that word. They did not, therefore, always provide a foundation for generalized, objective judgements on a particular linguistic or lexicographical point, but rather offered subsidiary evidence.

During the preparation of this edition, there was a general increase in interest in new words among lexicographers and their readers, particularly in North America and Great Britain. The two Barnhart Dictionaries of New English and their quarterly Companion presented well-researched material on neologisms. The Merriam-Webster Dictionaries published their record of new vocabulary in book form. Advance publicity for new dictionaries appearing throughout the world provided lists of the neologisms included in them. These were all valuable pointers to items that need to be included in the *OED*. And as ever, personal observation by members of the Dictionary Department and other contributors (the historical dictionary's own style of 'oral evidence') was a prolific source of further suggestions.

The selection of new words now included in the Dictionary is not intended to represent simply the emergent vocabulary of the last few years. To have selected solely according to this criterion would have highlighted terms which might not have achieved an established footing in the language, at the expense of other, older words. Indeed, one of the engaging aspects of historical lexicography is that terms which are held to be of recent origin often turn out to have existed for many years, though without achieving general currency. Recent neologisms have not been excluded, but have had to compete for their places alongside other candidates. In general, terms from most of the varieties of English are added, though the majority are current in American or British English, or in both. The vocabularies of modern computer technology, medicine, politics, economics, and popular culture are well represented, though not at the cost of terms from less obtrusive realms. Omission should not be equated with exclusion; a simple practical limit had to be set in accordance with the number of items that could be prepared in time for this edition.

There is no major difference between the editorial policy of the *Supplement* and that followed in the preparation of the further new entries; they are no more than an extension of the *Supplement* on a smaller scale, integrated, like the latter, into the main text of the Dictionary. In so far as these new entries have been compiled in accordance with the overall editorial policy of the second edition, they naturally embody a certain number of differences from the style of the *Supplement*. The stylistic changes have been described above, pp. xii–xvi. It was the policy of the *Supplement* to include earlier, further, and later examples of words and senses already in the Dictionary. This provision, being quite distinct from the addition of new entries, was not continued after the completion of the *Supplement*, but was deferred until such time as the full revision of the Dictionary may begin.

STATISTICS

This edition of the *Oxford English Dictionary* contains about 290,500 main entries, or about 38,000 (15 per cent) more than in the first edition. The extent of the text, however, has grown from around 44 million words to around 59 million, which represents an increase of 34 per cent. There are 350 million printed characters in this work. In addition to the headwords of main entries, the Dictionary contains 157,000 combinations and derivatives in bold type, and 169,000 phrases and combinations in bold italic type, making a total of 616,500 word-forms. There are 137,000 pronunciations, 249,300 etymologies, 577,000 cross-references, and 2,412,400 illustrative quotations.