Not just a pretty face: the contribution of typography to lexicography

Conference or Workshop Item

Published Version

Creative Commons: Attribution-Noncommercial-Share Alike 3.0


It is advisable to refer to the publisher's version if you intend to cite from the work.
Published version at: http://www.euralex.org/elx_proceedings/Euralex2004/

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading
Not Just a Pretty Face: The Contribution of Typography to Lexicography

Paul Luna
Department of Typography & Graphic Communication
The University of Reading
2 Earley Gate, Whiteknights
PO Box 239
Reading RG6 6AU, UK
p.luna@reading.ac.uk

Abstract
Typography – the rational organization of visible language – is a major contributor to making effective dictionaries. Dictionaries are structure-rich, and therefore require a potentially complex typography. This paper, by a practising designer of dictionaries, considers the repertoire of typographic effects that are used for both navigation through a dictionary (macro-typography), and also for the differentiation of individual structural elements (micro-typography). It considers historical examples (Estienne’s dictionaries, Johnson’s Dictionary, the OED) as well as more recent case studies including the author’s own designs for the latest edition of the Shorter Oxford English Dictionary. The primacy of the user’s needs emerges as a key theme.

1. The typographic repertoire
Book design is a utilitarian activity. It is intended to serve the reader by making the structure of the author’s text clear in a visual form, and also by making the book pleasant to handle and durable. My approach to discussing the design of dictionaries is to consider how the surface appearance of the page typography relates to the underlying structure of the author’s text. So first of all it is necessary to briefly define this awkward term typography. It has been used to describe a particular printing process (letterpress); in many languages it still just means printing; and students think of it as the art of laying out printed matter. None of these definitions express the role of visual organization in making the meaning of written language clear. It is not simply a matter of the arrangement of blocks of words, or the choice of particular typefaces to set them in, but an interactive process where rational visual choices are determined by inherent aspects of the author’s text, the expected reader, the kind of book being produced, and the production technology. If this stresses the collaborative nature of typography, then that is a good thing. Typographic design is something that grows out of a series of decisions about a text, and good typography can rarely be painted on to a text after the event. So basing its terminology on that of Twyman (1979), the working definition of typography for this paper will be ‘the rational organization of verbal graphic language’.

Tools for typographic analysis
One way of looking critically at a piece of design is to consider the genre it belongs to. Waller (1987) has given us tools for discussing typographic genre, and sees a genre as the interplay between the particular requirements of an author’s text (the ‘topic structure’), the
physical format (the ‘artefact structure’), and the means by which readers gain access to the information (the ‘access structure’). This approach has the advantage of placing the reader clearly in the frame. Typographers also consider a design by looking at macro-typographic issues and micro-typographic issues. These can be categorized at those which chunk and arrange information at a level for general access (macro), and those which provide close-grain differentiation (or harmony) of elements within those larger structures (micro). Aspects of macro-typography tend to relate to document and page layout, and micro-typography to what happens within a paragraph or within a line. A further tool for typographic analysis is to consider the configuration of written language: continuous prose, headings, lists, tables, algorithms, etc. (Twyman, 1979). Recognizing typographic configurations is something we all learn at an early age, and typographic genres gain much of their effect by exploiting this recognition of configurations.

Using Waller’s model, we can describe the topic structure for dictionaries as generally an alphabetic sequence of entries with further levels of information nested within them. The configuration that this involves is usually a list structure with various nested sub-lists. The artefact structure is generally a relatively compact book with a multi-column structure. The access structure is provided at the macro-typographic level by the alphabetic sequence and page elements (alphabetic sections, headlines, headwords) which assist overall navigation and the location of entries, and at the micro-typographic level by the differentiation of elements within an entry.

Dictionaries are by their nature structure-rich, and therefore require a potentially complex typography. Hanks (2003) has described dictionaries as ‘more highly structured than any other piece of text’. Because dictionaries have developed a repertory of typographic presentations, it is possible to describe how these map on to commonly-found structural elements, at least at in general terms. Svensén (1993) and Jackson (2002) both list the font variants currently considered normal for standard elements. Advertising provides evidence that dictionaries constitute a clear typographic genre; it calls on the typographic forms associated with dictionaries to imply an authoritative standard for a product. A recent UK press advertisement for a Honda car ‘defined’ the vehicle’s qualities, with ‘headwords’, ‘phonetics’, and ‘sense numbers’ set in typefaces similar to those used in Oxford dictionaries.

So what are the resources a typographer can call on? He can vary the font for a particular element, so that it contrasts more or less with its surroundings, he can alter sizes, and sometimes colours. These features can be described as ‘intrinsic’. ‘Extrinsic’ features that can be used are mainly to do with the deployment of space: entries can be divided into sub-paragraphs, various forms of indentation can be used (Twyman, 1986). While a hierarchical relationship between elements can be expressed by either intrinsic features or extrinsic features alone, a combination of features is the most normal approach. Each typographic effect is reinforced by another. This reflects the important role of redundancy in written language, which has been usefully illustrated by Garland (1966: 16) [Figure 1].

Allocating a typographic format (such as a particular font variant) to give emphasis to a particular word is described by Moxon, in his seventeenth-century manual of printing:
And what Words of Emphasis come in that precedent Matter; that he [the compositor] may Set them either in Capitals, Roman, Italick, or English [black letter]. (Moxon, 1683/4: §22 ¶5)

But Moxon's seems to be describing a rather ad-hoc process, and he is after all only discussing the preliminary pages of a book. The systematically mapping of typographic formats to text features, and thereby to structural elements, to differentiate them for the reader is now so normal, enforced by the rules of house style manuals (Walker, 2001), that it may need to be illustrated by an eccentric exception. It is now a generally accepted convention to distinguish book titles by italicization, so we are taken aback when we read George Bernard Shaw railing against the advice given by an early edition of Hart's Rules:

This is deplorable. To the good printer the occurrence of two different founts on the same page is at best an unavoidable evil. ... Not only should titles not be printed in italic; but the customary ugly and unnecessary inverted commas should be abolished. Let me give a specimen. 1. I was reading The Merchant of Venice. 2. I was reading “The Merchant of Venice”. 3. I was reading The Merchant of Venice. The man who cannot see that No. 1 is the best looking as well as the sufficient and sensible form, should print or write nothing but advertisements of lost dogs or ironmongers' catalogues: literature is not for him to meddle with. (Shaw 1902)

Fortunately Hart's prevailed in this matter. Shaw, of course, was writing at a time when an even-textured page (promoted by William Morris) was becoming the desired end of typography, and when it was felt that this could really only be achieved by returning to the closest English printing had to 'original' type, namely Caslon. Caslon, true to its eighteenth-century origins, had only a single weight. By using Caslon, you achieved uniformity of texture, because you could not use bold; roman, italic, and small capital fonts existed, but not bold. In general, dictionary typography has progressed towards more differentiation, not less.

2. Looking at the past

Approaching the history of dictionaries from the standpoint of typographic design is not a well-trodden path. (Hancher (1996, 1998), who considers the graphic and cultural qualities of nineteenth-century dictionaries, is an exception.) Perhaps this is because of the utilitarian nature of dictionaries, which are books with active readers who have a problem to solve, rather than contemplative readers who have an interest in the aesthetics of the book as an object. Typography develops over time, and the range of effects available in the earlier centuries differed from that available today. So the questions a typographer might pose of a dictionary are: to what degree does the typographic surface map the underlying structure in a way that is logical, consistent, and revealing to the reader? To what extent does it use (or even expand) the repertory of typographic effects for this kind of text?

Robert Estienne

The foundations of dictionary typography were laid by the Parisian scholar-printer Robert Estienne (1498–1559), who printed the Thesaurus Linguæ Latinae (1531) and the Dictionnaire François–Latin (1539) [Figure 2]. Estienne's spatial organization of the page
was exemplary, precisely because he mapped visual appearance to structural significance: he differentiated headwords by setting them on separate lines, in a larger point-size, indented; definitions were set full out, and examples of usage indented one em. Significantly, he introduced the differentiation of language by typeface, using italic type for French words and roman type for Latin ones (Twyman 1982). In other words, he used paragraphing and changes in type size to assist navigation, and changes of typeface to indicate structural elements within a paragraph. Estienne's preface to the Thesaurus includes the following comment on its design, which holds good for the Dictionnaire:

Here the different senses of words have not been indiscriminately piled into one confused heap, but distinguished by making each one start on a new line throughout, even in cases of twenty or more variant meanings. Here, as a result of this elegant differentiation in sequential form, the variety not only of constructions but also of expressions has been noted, by the observation of a regular manner of expression in authoritative writers. (Brandon, 1904: 42 n3, passage translated by Leofranc Holford-Strevens)

Estienne’s achievement looks even more significant when compared to the edition of Ortus Vocabularum (1509) printed in London by Pynson some thirty years earlier, which has only one typeface for both Latin and English text, and no indentation to articulate entries [Figure 3]. Subsequent printed English dictionaries were slow to aspire to such accessible typography, and only gradually adopted different typefaces and indentation systems.

Samuel Johnson

Relatively undeveloped typography might have been suitable when English dictionaries were little more than lists of hard words, but the development of lexicographical practice in the eighteenth century required a more considered approach. Bailey’s Dictionarium Britannicum (1730) entries contained elements and typographic presentations familiar today, such as etymologies set off in square brackets and cited forms in italic [Figure 4]. Martin’s Lingua Britannia Reformata (1749) introduced numbered senses, and the appearance of his page was surprisingly modern [Figure 5]. But it was Johnson who added significantly to the complexity by using not only numbered senses but also illustrative quotations. Johnson’s printer had limited typographic resources: the Caslon-like typeface that he used had only roman, italic, and roman small cap variant fonts. A comparison of folio, octavo, and quarto editions of Johnson’s Dictionary shows how much more effective it is to subdivide entries into paragraphs for senses and quotations, rather than to attempt to run elements on [Figures 6–9]. On the folio page (1755), the status of the illustrative quotations is demonstrated by their spatial arrangement, and the ranged-right italic sources provide an alternative means of access to the entries. On the quarto page (1773), senses are run on, and quotations are divorced from the senses they illustrate. The abridged London octavos (1758a, 1760) manage to provide some articulation through paragraphing; the Dublin octavo (1758b) does not.

James Murray and the OED

Mastery of a complex page with many fonts and paragraped senses was achieved by Murray in the OED [Figure 10]. The OED used a particularly effective combination of typefaces. The underlying face was an Old Style, effectively a much-regularized version of
Caslon. Its even and relatively light overall density on the page mean that two weights of bold-face type could be used to contrast with it. Murray also had a more complex system of illustrative quotations to contend with than Johnson—sequences of quotations demonstrate the changing use of words. The individually paragraphed quotation was not an option. The solution was to create banks of run-on quotations after each numbered sense. Bold dates were used to begin each quotation, providing both a visual catch for the start of each, and reinforcing the historical sequence of quotations.

The design of the \textit{OED} depends on the variety of types used, and is almost an exact counterpoint to Johnson's lack of variety. Murray's particular innovation in the use of bold was to use variant typefaces systematically to identify different structural items. Bold-face types for headwords in English dictionaries had appeared in Hunter's \textit{Encyclopædic Dictionary} in 1879. Murray used four bold founts: a large-face bold Egyptian, a large-face lighter Egyptian, a small-face lighter Egyptian—all cast on brevier (about $7\frac{1}{2}$ pt) and a nonpareil bold (about 6 pt). The boldest, largest fount was used only for headwords. The lighter large-face Egyptian was used for subsidiary headwords, usually archaic forms, and also for the sense-numbering system. The lighter small-face Egyptian was used within entries for variant forms, and the nonpareil bold was used only for quotation dates.

3. Focusing on the user

How can a dictionary be made more user-friendly? The move towards greater contextual information in dictionaries is an important trend. The development of corpus-based dictionaries has begun to emphasize the company a word keeps. As Hanks reminds us, 'innovation has been very much the exception rather than the rule in lexicography' (2003: 61). He also points out that while conventional dictionaries offer a number of senses to define each word, little help has traditionally been given to the reader to enable them to assess which is the most likely meaning for their purposes. Confronted with a sequence of senses to wade through, a reader needs every bit of help in selecting the most relevant. School dictionaries in the UK are required to provide word class, word history, and word family information (for pedagogical, not information-retrieval purposes), but there is no specification or requirement to indicate meaning through context (DfES, 2001). Carney (2004) argues that native-speaker dictionaries 'are about decoding, so naturally they privilege meaning over usage, semantics over syntax', and therefore lag behind bilingual dictionaries in the classification of phrases, and in definitions that respect the true syntactic and collocational boundaries of words.

Access devices

Recent dictionaries have experimented with access devices to help readers identify the most appropriate entry or sense [Figures 11–13]. Two approaches have been tested by Tono (1997): 'guide words' (used in \textit{Cambridge International Dictionary of English}, 1995) and 'signposts' (used in \textit{The Longman Dictionary of Contemporary English}, 1995). Matters are complicated because \textit{CIDE} follows the rule 'one entry, one core meaning', while \textit{LDCE} entries are subdivided into word classes. Tono concludes that the general approach of what he calls 'meaning-access indexes' is effective. Two further examples of this approach are the \textit{Macmillan English Dictionary} (2002), which uses an ruled-off 'menu' of senses at the start.
of each entry, and the Encarta dictionaries (1999, 2001), which start each sense with a ‘quick definition’ – a summary phrase in bold. The typography of these access devices could certainly be improved, and there has been a development towards more considered presentations in the Encarta Concise and MED. Recent work at The University of Reading shows that there is scope for developing the typography of meaning-access devices.

Typographic coding
As we have seen in the historical examples, spatial arrangement is probably more important than font differentiation in revealing structure. But this is problematic when dictionaries need to be compact and portable, or include more and more features for pedagogic or marketing reasons. When space is at a premium there is a tendency to use the less efficient differentiation by font change alone. The first editions of the Concise Oxford or the Pocket Oxford were miracles of compression, but hardly offered much assistance to the reader in navigating entries. Dictionaries that were typeset by hand had the advantage of flexibility in typeface choice and combination – the OED itself shows this particularly well. Subsequent editions of the COD increased the use of coding within entries. As entries became more complex, especially though the use of cross-references, problems arose of deciding which structural feature a particular word should be mapped to typographically. An example from COD6 shows the difficulty that arose when a complex structure met the relatively inflexible range of typographic options offered by mechanical composition.

COD6 was typeset in Monotype Baskerville, using a 15 x 17 matrix-case which allowed seven alphabets in a single size to be used simultaneously; these would normally have been roman upper and lower case, italic upper and lower case, bold upper and lower case, and small capitals. Small capitals were only available in roman – there were no italic or bold small capitals. Problems arose when a word had two different statuses in the entry structure, for example if it were both part of a phrase and also a cross reference; or if it was a repeat of the headword and part of a phrase. In the example [Figure 14], the two phrases given are ‘die game’ and ‘as game as Ned Kelly’. COD6 conventions replace the repeated headword with a swung dash – arguably satisfactory if the rest of the phrase retains its typographical coding to reveal its status. But in the specimen setting, die in ‘die game’ is a cross-reference, so the cross-reference formatting (roman small capitals) overrides the phrase formatting (italics), and as a result neither of the two words appear in the font that indicates a phrase. The problem was clearly noted, because the published dictionary amended ‘die game’ by the ad-hoc replacement of the swung dash with the word in full.

Digital fonts and mapping surface to structure
The ability of digital composition systems to combine many typefaces, font variants, and sizes without penalty, and the development of extended and related typeface families (groups of related families have been called ‘tribes’ by the type designer Gerard Unger) has provided a more appropriate set of resources for necessarily complex dictionaries, such as the Shorter Oxford English Dictionary. The seven-alphabet restriction has long been overcome, and the typefaces used in SOED5 (2002) represent a larger palette [Figure 15]. OUP Swift and OUP Argo are related designs, with matching appearing sizes (similar sizes align optically) and a consistent approach to character shapes. OUP Swift is a seriffed family, OUP Argo a sanserif
family, and both have full ranges of constituent fonts; small capitals, for example, are available in both typeface families in all weights and style variants. Figures are available in both ranging (capital-size) and non-ranging (lower-case-size) in all fonts. While there are no IPA, Greek or Cyrillic characters (yet) in these families, their appearing size relates so closely to Times New Roman that this typeface can be used for all those characters.

The seriffed–sanserif contrast can be used to separate definitions (seriffed) from metalanguage (sanserif), providing an axis of contrast analogous to the ‘bold for look-up items, normal weight for text’ that readers are used to. A further axis of differentiation can be the degree of expansion or condensation used. Specially-drawn condensed and expanded typefaces exist, but any digital fonts can be condensed and expanded in very fine increments. While this is normally avoided in good typesetting, it may be necessary in a dictionary to increase the ‘dynamic range’ of a typeface, and to optically correct very small sizes by making them relatively wider. Adjusting the size of type microscopically can ensure that a particular number of words fit on a page, or compensate for the spelling-out in full of abbreviations.

Design possibilities
Publishers may see the typography as a way of making dictionaries more ‘feature rich’: typographic quirks appear to be growing, for example the ‘associative’ use of different typefaces to differentiate notes in the New Penguin English Dictionary (2000), or the over-emphatic information boxes in the Oxford School Dictionary (2002). The real advantage of digital typefaces that can be fine-tuned more easily is to allow designers to use a wider range of typographic effects by carefully mapping form to content, so that typographic surface faithfully (and informatively) represents structure. Font variation need not be restricted to emphasis, but can distinguish definitions from metalanguage. Ambiguities, such as the use of one italic font to represent several elements, can be avoided. A word can display two statuses, for example a phrase can be in bold italic, with an embedded cross-reference in bold italic small caps. Importantly, the ability to micro-tune sizes allows a return to explicit paragraphing within entries and the expansion of abbreviations with little loss of economy or appearing size.

References

Dictionaries
— 1758a, 1760. A Dictionary of the English Language. Octavo ed. London: J. Knapton etc.
Other works


'Yes,' he replied, 'now the operation has become both difficult and dangerous.'

Figure 1: The usefulness of redundancy in graphic language (after Garland)

Figure 2: Estienne, *Dictionnaire François–Latin* (1539)

Figure 3: Pynson, *Ortus Vocabulorum* (1509)

Figure 4: Bailey, *Dictionarium Britannicum* (1730)

Figure 5: Martin, *Lingua Britannica Reformata* (1749)

English Faculty Library, Oxford. XW1[1749]

CEREMONY, n. [ceremonies, Lat.] 1. Outward rites; external forms in religion. Bring her up to the high altar, that's all. Sheflp, *Cofar.*

If you find them deck'd with ceremony. Sheflp, *Cofar.*

2. Forms of civility. Meeting were bare without it. Sheflp, *Maneth.*


David,Bodleian Library, Oxford. Vet. A5 e.6606

Figure 9: Johnson, *Dictionary* (Dublin octavo 1758)

Bodleian Library, Oxford. Vet. A5 e.6606

CERTAIN. adj. [certius, Lati.] 1. Sure; indubitable; unquestionable; indefectible; that which cannot be questioned, or denied. 2. Resolved; determined. 3. Unwavering; pure, plain, doubt. 4. Unfeeling; which always produces the expected effect. 5. Content; never failing to be. 6. Regular; solid; flat. 7. In an indescribable sense, form; as, a certain man told me this.

(1.) Thinks things are certain among men, which cannot be denied, without oblivion and folly.

The man is easily certain of, whether their ideas be more or less general.

(2.) However I wish that you had not my list.

Ceremonies to undergo like dream of death.

Concert with thee. Milton's *Par. Lyc.*

(3.) The form before Alciphron prefers.

To make her certain of the fall event. Dryden, *A.D.*

(4.) I have often wished that I knew as certain a remedy for any other diftemper.

Mind.

(5.) Virtue that dives us ways.

Though these dangerous uncertain paths. Dryden, *Epistles.*

Who calls the council, shall a certain day?

Figure 7: Johnson, *Dictionary* (quarto 1777)

Bodleian Library, Oxford. Don. d.224

Brigandine, n. [from brigand, Fr.] (f. *brigand*.) 1. Body armour composed of iron rings or small thin iron plates, sewed upon canvas, linen, or leather, and covered over with similar materials. *Planche de Civ. Cost.*; orig. worn by foot-soldiers and at first in two layers, hence in early quotes, in plural or as pair of brigandines; less strictly perf.

= coat of mail, corslet. *See Brigand.*

*1* 1596 Eng. *Chron.* (Camden) 66 Armed in a pair of brigandines. 1566 Paston Lett. 99 I. 129. 1 day payt of Brigandines kevert with blew fowete and 30th male, with leghirnaytree, the lower of the gown and the brigandyns vol 8. 1569 Acta Dom. Comiti 123 (Jan.) *The said Sirch Morgan held the brigandynes coutente in the commondy.* 1568 *Udal,* etc. Erasm. *Par. Mark* Pref. 4 They have thry brigandynes, there spunders gynide. 1599 Lane. *Wells* II. 86 A payt of brigandines. 1600 *Garnard* *Avi Warre* 9 *The Hallwardern, who was armed either with Brigandine or Corselet,* 1601 *Hussar* *Arm.* Furtho the spleches, and put on the brigandines *[Wyclif habitudes; Covers, breastplates; *Vulg. trac.*]...

Figure 10: Murray, *OED*
draught /drət/; BrE draught /drʌft/ n, AmE (BrE) a cold drink served from a large container rather than on tap (BrE) 

COLD AIR a current of cold air flowing through a room: Shut the window - there's a draught in here!

GAME 4 draughts (plural BrE, a) a board game played by two people, each with 12 round pieces, on a board of 64 squares; (BrE) b the pieces used in a game of draughts 

SWALLOW the act of swallowing liquid, or the amount of liquid swallowed at one time: Mick took a long draught of water.

on draught especially BrE a beer that is on draught is served from a large container rather than on tap (BrE) 

MEDICINE a medicine that you drink: a sleeping draught!

FIRE the flow of air to a fire 

SHIP the depth of water needed by a ship so that it will not touch the bottom of the sea, a river etc.

Figure 15: Comparison of specimen and printed entries in COD6

Figure 12: Macmillan English Dictionary, showing a 'menu'

game2 a. Like a gamecock, spirited (out2 ~; as ~ as Ned Kelly, (Austral. colloq.) very brave); having the spirit or energy to do; ready for. (cf. game1 in obs. sense 'fighting spirit')

vastidity /vɑːstɪdɪtɪ/ n, rare. E17. [Irreg. var. of vastity, after wds in -idity.] Vastness.

Shakespeare, Mac. for M. A restraint. Though all the world's vastity you had. To a determined scope.

vector /ˈvektər/ n. E19. [L = carrier, traveller, rider, from vect- pa. ppl stem of vehere carry, convey: see veh-.] 1 Astron. = radius vector s.v. radius n. Only in 18. 2 a math. A quantity having direction as well as magnitude, denoted by a line drawn from its origin to its final position. Cf. SCALAR n. M19. ↑ B Math. An ordered set of two or more numbers (interpretable as the coordinates of a point); a matrix with one row or one column. Also, any element of a vector space. E20. C Aerodynamics. A course to be taken by an aircraft, or steered by a pilot. M20. d Computing. A sequence of consecutive locations in memory; a series of items occupying such a sequence and identified within it by means of one subscript: spec. one serving as the address to which a program must jump when interrupted, and supplied by the source of the interruption.

Figure 14: Comparison of specimen and printed entries in COD6

Figure 13: Encarta World Dictionary, showing 'signposts'

lovely /ˈloʊvli/ adj. (-lier, -liest) 1. Beautiful and pleasing beautiful and pleasing, especially in a harmonious way. 2. Desirably very enjoyable or pleasant; lovely; having the spirit or energy to do; ready for. (cf. game1 in obs. sense 'fighting spirit')

lovenaking /ˈloʊvənkiŋ/ n. 1. Sexual activity sexual activity between lovers, especially sexual intercourse.

Figure 11: Longman Dictionary of Contemporary English, showing 'signposts'

Figure 10: Comparison of COD6 (left) and SOED5, showing how adjustments to microtypography can allow intra-entry paragraphing and the expansion of abbreviations