

Lexicographic slips: gathering and organising contextual data for dictionary entries¹

BRUCE L. FRASER

Faculty of Classics, University of Cambridge
blf10@cam.ac.uk

Introduction

This paper discusses the process of including contextual information in ancient Greek dictionaries, with especial focus on the task of organising the textual source-material.

First, contextual meaning is distinguished from formal meaning. Then, a description is given of the two functions of textual material: as citations in dictionaries to illustrate meanings, and also as the source-material for identifying meanings and writing the definitions.

The digitisation of ancient Greek texts has opened the corpus to systematic word-searches, and has led to a renaissance of Greek lexicography, in which new dictionary projects have developed their own procedures for collecting attestations. However, in order to use the results as sources for writing entries, the material must also be organised semantically into lexicographic 'slips'.²

An account is given of how lemmatising software has been used to identify and store the original textual passages cited in the Liddell-Scott-Jones dictionary, creating a digital archive of slips. This has been combined with a collection of other attestations identified in the texts, to create a comprehensive library of source-material for a Greek-English dictionary now being written at Cambridge. A description is also given of how the new contextual information gathered in this way is being incorporated in the dictionary.³

¹ Thanks are offered to Professor Manuel Alexandre Júnior and the team of the *Dicionário de Grego-Português* for valuable discussions at the 2006 Lisbon Colloquium, and for the opportunity to present this paper.

² The *Oxford English Dictionary's* definition of this sense of the word 'slip' is given in footnote 11.

³ Further information on the *Cambridge Greek Lexicon* can be found at www.classics.cam.ac.uk/glp/

1. Two sources of meaning

1.2. *Word-form*

A major factor in choosing the wording of definitions is that, as Chadwick (1996: 21-23) points out, in all languages 'nearly all words belong to families', and this 'family resemblance' is central to the determination of meaning. It is especially important in ancient Greek, where word-compounding was very productive, and where affixes make a major contribution (see Buck and Petersen 1944).

This does not mean that we identify a word's meaning from its *etymology*, but rather from its *derivation*: that is, the transparent combination of morphemes that it contains (Zgusta 1987: 259). However, this is only the starting-point for the process of writing a dictionary entry, because it gives only a single meaning.

1.2. *Usage: textual evidence*

The major problem of lexicography, as of semantics, is that words generally do not have just one meaning, but a range of senses, which vary in time, genre, and their interactions with other words. The distribution of the senses will be reflected in the structures of dictionary entries, which are consequently organised, not just as a list, but as a hierarchy of sections and subsections, each of which corresponds to a meaning or sub-sense. The ensuing dictionary entry is therefore as much a *semantic map* (Adrados 1977: 265) as a description of changing meanings.

In Greek and English lexicography of the nineteenth and twentieth centuries, the most influential method of ordering the senses was to give them in chronological order, in association with examples taken from the literary texts. The first (modern) alphabetic Greek dictionary, and the first dictionary from Greek to a modern language, Schneider (1797-8), was also the first to give extensive citations to exemplify meaning, but these were taken mostly from early epic, and had little historical range. Passow (1831), in his *Handwörterbuch der griechischen Sprache*, used Schneider's citations as the basis of his own work, but increased the scope of the citations, in order to map what he called the life-story, the *Lebensgeschichte*, of each word. The same approach was adopted by Liddell and Scott (1843: v-vi), who cite his remark that he had 'found it necessary to go below Homer and beyond Schneider', and who declare that their own plan 'has been that marked out and begun by Passow, viz. to make each article a History of the usage of the word referred to.'

However, they continued, 'It will be understood that deviations from the strict Historical order must occur. Homer will sometimes use a word in a metaphorical sense only, the literal sense of which first occurs (perhaps) in Plato. In such instances, of course, we give the literal and actual sense the preference.'

This semantic modification was in time formalised as 'logical' ordering: an idealised version of the recorded sequence, which visualises changes of

meaning as an evolutionary process. The principle is described by Murray (1888: xxi): 'If... we possessed written examples of all the uses of each word from the beginning, the simple exhibition of these would display a rational or logical development. The historical record is not complete enough to do this, but it is usually sufficient to enable us to infer the actual order.'⁴

However, whatever criteria are chosen for *organising* the senses, all dictionary-writers still *identify* them by using the evidence of the literary texts. The information in a dictionary entry is therefore in great part dependent on the interpretation of the textual citations, which provide the main evidence for meanings. Most of the lexicographer's time is spent on identifying textual senses, and that is also the focus of this paper. The source-material can be categorised in two groups: previously-collated citations, and other attestations which have been discovered in the texts. In the next section of this paper, these two groups are discussed in turn.

2. Textual citations from TGL to TLG: a brief account

2.1. *Previously-collated attestations*

We firstly have available to us the citations gathered in previous dictionaries, most notably Liddell-Scott-Jones (LSJ) and its Supplements, the *Vocabolario della lingua greca* (GI), and the *Diccionario griego-español* (DGE) and the *Lexikon des frühgriechischen Epos* (LFgrE) now in progress.

These volumes stand in a tradition of continuous accumulation of exemplary textual material. In the four centuries from the publication of the *Thesaurus Graecae Linguae* of Stephanus in 1572 to the establishment of the *Thesaurus Linguae Graecae* research centre in 1972, citations have become a distinctive component of lexicon entries. In the work of Stephanus, where words were grouped by 'family resemblance', brief phrases were given as examples, without line or chapter references (although authors, and sometimes works, were cited). The later, alphabetic, editions of the *Thesaurus* (Valpy and Barker 1816-28, Hase 1831-65) introduced referenced citations, but these were very brief: often just one-word quotations from the early grammarians and lexicographers, rather than illustrations of usage.

The first (modern) alphabetic Greek dictionary, and the first dictionary from Greek to a modern language, Schneider (1797-8), used more extensive citations, mostly from early epic, as examples. These provided the core source-material for subsequent Greek lexica: Passow (1831) drew on them for his citations, and Liddell and Scott (1843) in turn used his material as the basis for their own.

In their seven subsequent editions, Liddell and Scott steadily increased the number and range of quotations, drawing on the alphabetic *Thesaurus* of Valpy and Barker, and then on a variety of later sources, as the discoveries

⁴ 'Logical' ordering is also described in HIORTH (1955), KIPFER (1984), ZGUSTA (1987, 1989, 2006), SILVA (2000), and FRASER (2008 forthcoming).

and textual editions of the nineteenth century unearthed new attestations, until the accretion of new material made a complete reworking necessary.⁵

In 1904, a proposal was made to the British Academy for the creation of a new Thesaurus, in order to organise the newly-discovered material.⁶ However, in a memorable phrase, which has frequently been cited, Hermann Diels (1905: 693) compared the task of collating the citations from the full corpus of ancient Greek literature as equivalent to 'in dieses Chaos den Nus hineinzubringen',⁷ and the task was eventually abandoned as unfeasible, in favour of a further revision of Liddell and Scott's lexicon, which was published in ten parts from 1925 to 1940 as its ninth edition, LSJ.

This great work has proved to be the foundation of subsequent Greek lexicography, but it may perhaps be described as a *magnificent failure*, because so much new material has been incorporated into the structure of the eighth edition that the clarity of the semantic descriptions is often overwhelmed: see Zgusta (1987, 271-2), Glare (1987) and Chadwick (1994). Since then, the ever-increasing volume of new material has been collected in independent volumes: new citations were published in Supplements to LSJ (1968, 1996), and the historical range was extended by Lampe (1961-8) and Trapp (1994-9).

2.2. Digital access to the texts

Diels' goal of bringing Νοῦς into the chaos appeared to be no nearer to fulfilment, until a fundamental technical breakthrough: the digitisation of almost the entire corpus of extant Greek texts, which allows dictionary-writers to make automated searches for every word-form. The main publisher of the literary texts is the *Thesaurus Linguae Graecae*, whose first CD-ROM was released in 1985. Since 1987, the Perseus digital library at Tufts University (www.perseus.tufts.edu) has been making available online a collection of Greek and Latin texts, with translations and a wealth of secondary material, and since 2001 the TLG has also been published online (www.tlg.uci.edu). A burgeoning number of research groups are now working on the digitising of inscriptions.⁸

A great range of software has been developed to search these digital libraries for individual word-forms, and this has coincided with a remarkable renaissance of Greek lexicography. The DGE is now combining the sources from LSJ and Lampe with the latest papyrological and epigraphic discoveries, and rethinking the semantic analyses of LSJ, while epic vocabulary is being

⁵ ZGUSTA (1987: 264-72) and GLARE (1987) give contrasting accounts of the changes in LIDDELL and SCOTT's approach. Their last (eighth) edition was published in 1897, the year of SCOTT's death and a year before LIDDELL's.

⁶ For a brief account of the discussions, see LSJ (1925: iv-vii).

⁷ 'Bringing Νοῦς into this Chaos.' The expression is cited in LSJ (1925: v), BERKOWITZ and SQUITIER (1990: vii), PANTELIA (2000: title).

⁸ Notably the Packard Humanities Institute (epigraphy.packhum.org/inscriptions), the Centre for the Study of Ancient Documents (www.csad.ox.ac.uk/CSAD), and the groups of the EpiDoc collaborative (epidoc.sourceforge.net).

re-examined in the LfgrE.⁹ The GI (1995) gives us a unified collation of citations from LSJ, Lampe, and new sources, in a highly-readable format, for use by the student as well as the scholar; also published in 2004 as a CD-ROM. Lexicons from Greek to other languages are now in progress, notably the *Woordenboek Grieks* at Leiden (www.woordenboekgrieks.leidenuniv.nl) and the *Dicionário de Grego-Português* at Lisbon (lexiconpt.no.sapo.pt).

The use of the TLG word-indexes now makes it feasible for lexicographers to find all the citations for each word-form appearing in the corpora. This is extremely helpful during the later stages of writing entries, for the task of checking senses and attestations, as described by Adrados and Somolinos (1994), and Somolinos and Berenguer (2005).

However, even with fast access to the digital collections, the task of organising the results is still very labour-intensive. Because ancient Greek was a highly-inflected language, with some lemmas having as many as a thousand forms (Crane 1991: 243), searching for all of them is still very time-consuming. Programs were therefore developed to collate the results of searches in selective ways, such as by authors and by individual works, as well as searches of the full corpus. A notable example is the Scuola Normale Superiore (SNS) software (smsgreek.sns.it/sns.html) developed by the Istituto di Linguistica Computazionale at Pisa, and used by the GI and the DGE.¹⁰

The key advance in search technology was made with the development of the Perseus morphological analyser (henceforth *Morpheus*), which can recognise inflected forms as belonging to a particular lemma: see Crane (1991). This opened the whole digital corpus to systematic lexicographic study, by unifying the search process itself: we can search for every attestation of a lemma, rather than only for individual forms. This means that we can quickly find all the attestations needed for each dictionary entry.

However, we should remember that, despite its name, the TLG is not a thesaurus, in the tradition of Stephanus, but a digital library: a collection of texts organised by authors and works. To create a thesaurus from it, we need to collate the word-forms by meaning. *Morpheus* gives us the forms we need, but we still need to sort them semantically. This would be a very slow process, if we had to start from scratch. We need a way to look again at the sorting which has already been done.

3. Lexicographic slips

3.1. *Identifying senses: the past*

The lexicographic purpose of a quotation is not only to identify an attestation from a particular author, but also to arrive at an interpretation of it. In order to do this, many citations need to be compared as each entry is composed, and they must first be organised into semantic groups.

⁹ The semantic approach of the DGE is described by ADRADOS (1977, 1986, 1997, 2000).

¹⁰ See ADRADOS and SOMOLINOS (1994: 244-5).

In traditional lexicography, the manual collection of the source material constituted a part of the semantic sorting. The *Oxford English Dictionary* (OED) provides the paradigmatic example: the team of volunteer readers who collected citations were required to make judgments on meaning, as well as noting the occurrences of words. Murray (1879: 5) asked them to 'Make a quotation for *every* word that strikes you as rare, obsolete, old-fashioned, new, peculiar, or used in a peculiar way', and to 'Make *as many* quotations *as convenient to you* for ordinary words, when these are used significantly, and help by the context to explain their own meaning, or show their use.' Automatic word-searches cannot, of course, perform this function.

The second and principal part of the sorting process was performed by the OED editors, using physical manipulation. Murray (1884: 510-511) describes the laborious process by which the paper 'slips' (half-sheets of note-paper, each bearing a single citation) were arranged by the editors to create semantic maps for each word.¹¹

'Only those who have made the experiment', he wrote, 'know the bewilderment with which editor or sub-editor, after he has apportioned the quotations... and furnished them with a provisional definition, spreads them out on a table or on the floor where he can obtain a general survey of the whole... shifting them about like pieces on a chess-board, striving to find in the fragmentary evidence of an incomplete historical record, such a sequence of meanings as may form a logical chain of development.'¹²

3.2. *Identifying senses: the task*

Modern lexicographers need to carry out a similar sorting process, but we do not have the benefit of pre-existing paper slips. As described above in 2.1, the citations used in LSJ were not derived from a collection of slips, but had been added cumulatively in the eight previous editions of the lexicon.

Consequently, the only record of the textual sources is constituted by the references and brief quotations in LSJ itself. Because, as Glare (1987: 17) notes, these quotations are too short to confirm meanings, we cannot assess and revise the semantic groupings and definitions without going back to the original passages. We must, as it were, reconstruct the hypothetical slips-archive on which the entries had been based. Then, after studying this selection of citations, we need to examine further attestations, to search for other (and especially chronologically-later) meanings, and use them to re-organise the semantic map for each lemma, according to the full evidence.

¹¹ Sense 10 of the noun 'slip', homonym 2, is defined in OED as a 'piece of paper or parchment, esp. one which is narrow in proportion to its length.' Its lexicographic use is illustrated in the dictionary with two citations: as a noun, from 1846 (F. MADDEN, *Lazamon* I, Preface xli), 'after writing near 50,000 slips, it was found impracticable to carry the design [of the glossary] into execution', and as a verb (homonym 4, sense 2), from 1902 (*Athenæum* 23 August 256.1), 'Miss Betham-Edwards's new story .. is being 'slipped' by Dr. Wright .. for his Dialect Dictionary.'

¹² See SILVA (2000: 89-90) for further description of the methodology.

4. Identifying senses: electronic slips

These considerations were the stimulus for the development of a digital resource for the Cambridge dictionary, to build on the possibilities of digital searching. We wished to re-examine all the citations in LSJ, and we also needed a fast method of examining other citations which had not been included in the dictionary. The possibility of using the *Morpheus* software for these tasks was the inspiration for a project initiated in 2000 by A. A. Thompson of Cambridge University and J. A. Rydberg-Cox of the Perseus Project, and developed over the following five years by Rydberg-Cox and B. L. Fraser. The result is a two-part HTML archive, linked together in a unified structure, which can be easily used by editors who are more accustomed to consulting books in a traditional library. In order to give the necessary speed of access to the citations, the archive is built from the results of hundreds of hours of systematic digital searching.

4.1. Pre-searching

Because we can predict every TLG search that we will eventually want to perform, a program was designed by Professor Rydberg-Cox to conduct these searches in advance. Our corpus of texts was pre-searched for all word-forms, using *Morpheus* to group them by lemma; and then every attestation, together with a passage of surrounding text, was archived in static HTML pages. This provides the dictionary writers with immediate access to the searches, and also enables the citations and their contexts to be available in a generic format, rather than being tied to any specific operating system or database program.

To enhance speed of access for the users, the program matches the Greek texts to the English translations which are available on Perseus, and, for cross-platform compatibility, the collection is duplicated using Super-Greek and Unicode fonts.

4.2. Separating the LSJ citations

The larger part of the archive is composed of citations arranged by authors in chronological order, for a canon of 70 writers from Homer to Plutarch. In a separate operation, the analyser collects all the attestations from the corpus which are cited in the LSJ entry for that lemma, using the digital edition of the dictionary available on Perseus. This semantically-organised collection of textual passages constitutes a retrospectively-created realisation of the hypothetical LSJ slips-archive which was mentioned above in 3.2.

Both collections are stored in a combined archive, using 20 gigabytes of disk space: equivalent to thirty CD-ROMs, and so much greater than the total size of the digital corpus. (It has a high level of redundancy because each sentence must be repeated for every word-form that it contains.)

Because the listing of citations on each HTML page can be very long, only the first two citations are shown in this illustration: Herodotus 6.67 and Thucydides 8.93, the first citations in the LSJ entry. However, all the citations appear on this one page, in the same order as they appear in LSJ. (The other four pages of citations are those which do not appear in LSJ, and are organised by author, in chronological order.)

6. Key features of the Cambridge slips

The archive has five key features which make it a highly-effective lexicographic tool: (1) the separate collection of citations from LSJ, (2) a provision for checking ambiguous lemma-forms, (3) provision for checking missed LSJ citations, (4) automatic citation matching between textual editions, and (5) a unified display of citations collected from three collections of texts. These are described next.

6.1. The LSJ collection: the 'weave'

In order to make maximum use of the semantic sorting which has already been performed on the LSJ citations, we also display them in what we call a 'weave': that is, interwoven with the text of the LSJ entry itself. The start of the weave display for the same word as shown in Figure 1, θέατρον, is shown in Figure 2:

Previous Entry θεατροπόρος	Next Entry θεατροποιός	Greek and English Slip or Greek Only Slip	Unwoven Entry	Citation Summary	Table of Contents
---	---	---	-------------------------------	----------------------------------	-----------------------------------

θέατρον, Ἴον θεητρον, τό, (θεαμαί)

A place for seeing, esp. for dramatic representation, theatre, Hdt 6.67

— 6.67 ἀνέστη δὲ τὸ ἐκατομμύριον αὐτὸ δὲς αὐτὸς αὐτὸν ἀποκρίσας ὅτι περὶ τῆς θέας, αὐτοὺς δὲ οἱ, τὴν αὐτὴν ἀποκρίσας ταύτης δόξα Λακεδαιμονίᾳ ἢ αὐτῆς ἐκείνου ἢ αὐτῆς ἐκείνου, ταῦτα δὲ αἶμας καὶ κατακλιθεὶς ἵκεν ἐπὶ τοῦ θεάτρου ἐπὶ τὸ εὐνοῦν αὐτῶν, αὐτῶν δὲ μεσοσταθίου ὅτι τὸ αὐτὸ ἴσως δὲ τὴν αὐτὴν ἀπόδοσιν ἀποκρίσας δὲ τὴν αὐτὴν δόξαν δὲ τὰς αἰσῶν οἱ τὸν ἐκείνου κατακρίσει, ταῦτα λέγειν. — He was grieved by the question and said that he had experience of both, while Leotychides did not, and that this question would be the beginning for Sparta of either immense evil or immense good fortune. He said this, covered his head, left the theater, and went home, where he immediately made preparations and sacrificed an ox to Zeus. Then he summoned his mother.

, IG22 1176, al.; as a place of assembly, Th 8.93

— [1] τῇ δ' ἰσοταρίᾳ αἱ μὲν τετρακονταὶ ἐς τὸ Βουλευτήριον θύειν καὶ τετρασθῆναι (ἐκείνοισι οἱ δ' ἐν τῇ Πειραιεὶ ὄντι τὸν τετράκοντα ὄν) ἐκείνου δὲς αὐτὸς αὐτὸν ἀποκρίσας ὅτι περὶ τῆς θέας, αὐτοὺς δὲ οἱ, τὴν αὐτὴν ἀποκρίσας ταύτης δόξα Λακεδαιμονίᾳ ἢ αὐτῆς ἐκείνου ἢ αὐτῆς ἐκείνου, ταῦτα δὲ αἶμας καὶ κατακλιθεὶς ἵκεν ἐπὶ τοῦ θεάτρου ἐπὶ τὸ εὐνοῦν αὐτῶν, αὐτῶν δὲ μεσοσταθίου ὅτι τὸ αὐτὸ ἴσως δὲ τὴν αὐτὴν ἀπόδοσιν ἀποκρίσας δὲ τὴν αὐτὴν δόξαν δὲ τὰς αἰσῶν οἱ τὸν ἐκείνου κατακρίσει, ταῦτα λέγειν. — He was grieved by the question and said that he had experience of both, while Leotychides did not, and that this question would be the beginning for Sparta of either immense evil or immense good fortune. He said this, covered his head, left the theater, and went home, where he immediately made preparations and sacrificed an ox to Zeus. Then he summoned his mother.

, Lys 13.32

— [31] μετὰ τοῦτο ποσειδάωνος ἀνέστης τὴν πόλιν τὴν [32] ἐκείνου δὲς αὐτὸς αὐτὸν ἀποκρίσας ὅτι περὶ τῆς θέας, αὐτοὺς δὲ οἱ, τὴν αὐτὴν ἀποκρίσας ταύτης δόξα Λακεδαιμονίᾳ ἢ αὐτῆς ἐκείνου ἢ αὐτῆς ἐκείνου, ταῦτα δὲ αἶμας καὶ κατακλιθεὶς ἵκεν ἐπὶ τοῦ θεάτρου ἐπὶ τὸ εὐνοῦν αὐτῶν, αὐτῶν δὲ μεσοσταθίου ὅτι τὸ αὐτὸ ἴσως δὲ τὴν αὐτὴν ἀπόδοσιν ἀποκρίσας δὲ τὴν αὐτὴν δόξαν δὲ τὰς αἰσῶν οἱ τὸν ἐκείνου κατακρίσει, ταῦτα λέγειν. — He was grieved by the question and said that he had experience of both, while Leotychides did not, and that this question would be the beginning for Sparta of either immense evil or immense good fortune. He said this, covered his head, left the theater, and went home, where he immediately made preparations and sacrificed an ox to Zeus. Then he summoned his mother.

, SIG976.4 (Samos, ii BC), Posidon 36 J., Act. Ap. 19.20, etc.; θ. ἰσχυρῶς τὸν, of the Roman amphitheatre, D.C. 43.22; ἐπὶ τὸ θ. εἰσέρπειν to bring upon the stage, Isoc 12.122

— οἱ παῖδες θύειν ἐπὶ τὸν ἀποκρίσας ἀποκρίσας ὅτι περὶ τῆς θέας, αὐτοὺς δὲ οἱ, τὴν αὐτὴν ἀποκρίσας ταύτης δόξα Λακεδαιμονίᾳ ἢ αὐτῆς ἐκείνου ἢ αὐτῆς ἐκείνου, ταῦτα δὲ αἶμας καὶ κατακλιθεὶς ἵκεν ἐπὶ τοῦ θεάτρου ἐπὶ τὸ εὐνοῦν αὐτῶν, αὐτῶν δὲ μεσοσταθίου ὅτι τὸ αὐτὸ ἴσως δὲ τὴν αὐτὴν ἀπόδοσιν ἀποκρίσας δὲ τὴν αὐτὴν δόξαν δὲ τὰς αἰσῶν οἱ τὸν ἐκείνου κατακρίσει, ταῦτα λέγειν. — He was grieved by the question and said that he had experience of both, while Leotychides did not, and that this question would be the beginning for Sparta of either immense evil or immense good fortune. He said this, covered his head, left the theater, and went home, where he immediately made preparations and sacrificed an ox to Zeus. Then he summoned his mother.

This display is more informative than the 'list' format illustrated in Figure 1, in two ways. Firstly, it gives us a check on accuracy: we can easily see whether any citations are missing. Secondly, it gives us semantic information: we can see the LSJ definitions next to each passage, and so compare their interpretations with ours. Because the citations are given in the order of the semantic groups of LSJ, we can benefit from the semantic sorting which has already been done, and make it the reference-point for our own revision. Three senses are visible here: the basic meaning of theatre as a place for dramatic performances (Herodotus), its use for political meetings (Thucydides), and a more abstract sense, *the stage, the theatre*, referring to the representations (Isocrates). The illustration does not show the full HTML page, which includes a fourth, collective, sense, *spectators, audience*.¹³

6.2. Checking ambiguous lemma-forms

Of course, automatic searching has its limitations, and there will be instances where the program fails to recognise the passage corresponding to a LSJ citation, or the correct lemma from which an inflectional form is derived. If such failure leads to serious loss of time, then the archive will be, in practical terms, of limited value. In order for it to be a usable research tool, we need to have facilities to cope immediately with the failures.

The most common problem is failure of lemma-identification. This has two possible causes. Firstly, the morphological analyser cannot identify every word-form. It is limited by the size of its index, which includes about 97,000 Greek stems and 14,000 inflections. This enables it to recognise 69% of the word-forms in the Perseus texts (constituting about 99% of the attestations). That gives a level of accuracy of about 85%: a good percentage, but still resulting in a substantial number of unresolved forms and missed citations.

The second possible cause of failure is that the process of lemmatisation is itself fundamentally limited by the presence of ambiguous forms: ἄννα, for example, could be the vocative of ἄνναξ, the Aeolic feminine of ἄνη, or the anastrophic form of ἀνά (or perhaps even a neuter plural of ἄνοος). However, we find that, in practice, homonyms like ἄννα or λῆξις cause least difficulty, and complexities of verb inflection cause most.

To meet these eventualities, the program is therefore designed to give us automatic feedback, by identifying the level of certainty in lemma-identification, and assigning a 'weight', or probability-number, to each attestation, which is based on the number of possible lemmas from which the form could be derived (as far as the program recognises). This is the basis for the totals of 'unambiguous' and 'ambiguous' citations shown in Figure 1.

The ambiguous forms must then be lemmatised manually. In practice, this does not take long: the eye can very quickly scan down a page of chrono-

¹³ The non-LSJ slips show that the two concrete senses appear throughout Greek, while the abstract sense is much less common. The development of the collective sense is especially interesting, being the usual sense in Aristophanes and in Plato, who gives it a much more general application, to any kind of audience or group of spectators. A fifth sense, *what is seen, spectacle*, is not identified in LSJ, but appears in the New Testament.

logically-arranged citations. We had originally intended to use our own work-results to improve the accuracy of the program, and to enter our corrections back into the *Morpheus* index using an online user-form, but, unfortunately, time constraints on the project precluded the development of this facility.

Eventually, the ideal solution would be to submit the output of the lemmatisation to a syntactic parser, which could disambiguate many of the remaining uncertainties. Probabilistic methods of syntactic parsing have been developed for English by the COBUILD team (Karlsson et al. 1995), but analysis in highly-inflected languages with variable word order depends how the digital text has been encoded in the first place. If all the words in a text were tagged individually by part-of-speech, they could be parsed to create a 'tree-bank'. Work on parsing Latin texts is now under way, and, in the mid-term, we may expect parsed Greek texts to become available.¹⁴

However, we needed to use the archive immediately, and so we required a strategy to cope with identification failures. Our solution was to combine the feedback with text-links. Every failure-report is accompanied by a hyperlink to the passage which was searched, so that we can check the text, by clicking on the link. The small horizontal lines preceding all the text passages shown in Figures 1 and 2 are the hyperlinks. We have, as it were, embedded the slips archive within the digital library of texts. This allows us to check problems immediately, reducing the times when we have to leave our workstations and consult the print editions.

6.3. Checking missed LSJ citations

A similar procedure is used for failed identification of LSJ citations. The program indicates to us where it has failed to find the word-form in the cited passage, and we can then immediately check the text. This feature can be

Previous Entry ἀβακτηούς	Next Entry ἀβακτητον	Greek and English Slip or Greek Only Slip	Unwoven Entry	Citation Summary	Table of Contents
---	---	---	-------------------------------	----------------------------------	-----------------------------------

ἀβακτης , ἐς , (βαζω)

A: *speechless: hence, calm, gentle*, ἀβάκην (Aeol. acc.) τὰν φρέν' ἔχω
 Sapph.72. Adv. κέως, εὐδοίῃσι Poet. ap. EM2.57: also ἀβακῆμαι Hsch., ἄβαξ
 Lex. Rhet. ap. Eust.1494.64.

There are 1 citations from the IGL corpus in the electronic LSJ entry for ἀβακτης .
 The program located 0 citations and failed to locate matching text for 1 citations.

The program thinks it failed to identify the following:
[Sappho 72](#)

¹⁴ For discussion on parsed texts, see ABEILLÉ (2003), and for Classical languages BAMMAN and CRANE (2007).

illustrated for the word ἀβάκης, *speechless, calm*, whose LSJ entry is shown in Figure 3.

We can see from the absence of an inserted passage that *Morpheus* has missed Sappho fragment 72, and the feedback at the bottom of the page confirms this. By clicking on the hyperlink, the underlined ‘Sappho 72’, we move directly to the fragment, which is shown in Figure 4:

The screenshot shows a web interface with a 'Table of Contents' link on the left and a 'Go to' field on the right containing the text '1.72. LP 120.1'. Below these elements, a line of Greek text is displayed: 'ἀλλά τις οὐκ ἔμμι παλιγκότῳι ὄργανι. ἀλλ' ἀβάκην τὰν φρέϊ' ἔχω ...'. In this text, the words 'ἀλλά', 'οὐκ', 'ἔμμι', 'παλιγκότῳι', 'ὄργανι', 'ἀλλ'', 'ἀβάκην', 'τὰν', 'φρέϊ'', and 'ἔχω' are all underlined, indicating they have been identified by the analyser.

In this fragment, the words which the analyser has identified are all underlined as parsed, and we can see that ἀβάκην is in fact there, but unrecognised (because it is a paroxytone accusative form not listed in the *Morpheus* index). So we still have fast access to the correct citation, even when the program has failed to identify the form. The consequent saving in time is substantial: this feature transforms the slips database from an ancillary tool with excellent but limited coverage, into a dependable, ‘all-weather’ reference system.

6.4. Citation matching

In order to identify all the LSJ citations, we also need to match any variations in numbering. In general, the citation systems for Greek texts are remarkably stable: the LSJ line numbers for Homer and the tragedians, and the section numbers for the prose texts, are much the same in modern editions. However, the texts of many early poets, especially the lyricists, have been republished in new editions which give different fragment numbers. We have therefore compiled a concordance from LSJ to the modern editions of the lyric and iambic poets, and also to epic, comic, and tragic fragments, where modern editions differ from LSJ.

This ‘poetry map’ is integrated in the electronic database. Its use can be demonstrated from the citation from Sappho shown in Figure 4 above. LSJ cites this as fragment 72 in Bergk’s *Poetae Lyrici Graeci*, while TLG uses Lobel-Page’s *Poetarum Lesbiorum Fragmenta*, where it is fragment 120. By tagging it with the LSJ number, and also mapping that to the modern edition number, we can ensure that the LSJ citation is always recognised, even in cases like this where *Morpheus* fails to find the target-word.

6.5. *Unified display of citations from three separate text collections*

The archive has another unique feature: its triple combination of Greek texts from the Perseus and the TLG collections, plus English translations from the Perseus library. This not only represents a considerable technical feat of data-mining, but is also an impressive example of collaboration between Perseus and TLG, who gave permission for their texts to be used. The core of the archive is the Perseus library, which provides the Greek texts which have stable textual readings. When each word-form is identified and the chunk of surrounding text selected, that specific sentence-file in Perseus is matched to the corresponding sentence-file in the English text. This enables a matching passage of English text to be displayed below each Greek one, helping the lexicographers to scan quickly through the texts.

For those authors and works which are not stored in Perseus, the TLG texts are mined, and the resulting quotations are given in their correct positions in the display. The first indication that we are looking at a TLG text will usually just be the lack of a English translation. This seamless transition between the Perseus and TLG texts ensures that we have a complete coverage of our corpus texts.

6.6. *The slips: summary of lexicographic functions*

The archive gives us a digital library tailored to our needs, with exceptionally fast access, because it displays the results of millions of searches, with the words collated with their contexts and indexed for reference. A lexicographically-useful size of passage is selected, set at three sentences, which gives us enough context to evaluate the word meanings.

The database is proving indispensable in the writing of our lexicon articles, and has transformed the nature of the project, by allowing us to examine the texts as we write, and to compare the LSJ citations with the others. Pre-searching has proved to be a highly-effective way of utilising the limited time available for writing the dictionary. The HTML format is also very user-friendly: we can navigate very quickly between the two components of the double archive (the LSJ citations and the others). The failures of identification cause minimal problems, because, as described above in 6.2, every page of the archive is linked to the full texts. In sum, without this resource, it would have been impossible to write fresh definitions, unless we had a much larger team of writers and much more time.

7. The future

The slips archive could be used to provide the textual references for the digital edition of the lexicon. We intend to make it freely available for interested scholars, linked with our online edition, which will be published on the Perseus site. The poetry map is available online now.¹⁵

¹⁵ A 'human-readable' version of the poetry map is published online at www.chlt.org/lexicon/papers/Poetry_Map.pdf.

Since we developed this resource, lemmatisation based on the *Morpheus* technology has (from 2006) become available on the TLG website. This gives all researchers the ability to make integrated searches for every lemma-form, and, because it covers all the TLG texts, we use it ourselves, as an invaluable complement to our archive.

The architecture of the pre-searched archive appears to have general applicability, rather than being limited to this dictionary. Even more importantly, the LSJ weave allows us to study the textual passages underlying the LSJ citations. Because the database could easily be extended to cover more authors and texts, editors of future Greek lexicons may well wish to use the same method to create their own slips.

8. Integration of the new contextual data

The archive has two primary lexicographic functions: for checking meanings and attestations, and for identifying new ones. Consequently, it also enables the editors to conduct new research: because we review all the LSJ citations, and a good number of others, we have the opportunity to gather new contextual information as we write each entry.

For example, we can collate not only the syntactic constructions which depend on a particular verb, but also the kinds of subjects which govern it, and in our dictionary entries we match both these groups with the lemma meanings. Similarly, for adjectives, we list the classes of nouns which they qualify. For noun entries, we note words given in contrast to the lemma, and also connotations (such as approval or disapproval) which are implicit in the context.

Because the Cambridge dictionary is being written for students, it has two practical advantages in putting this information at the heart of its method. Firstly, as Greek quotations are rarely given, more space can be given to semantic description.¹⁶ Secondly, because the citations are restricted to literary authors, omitting inscriptions and the smaller fragmentary texts, meaning is usually identifiable from the context.

We may summarise the intra-linguistic context as *collocational* information, whose importance was summarised by Firth (1957: 179) in the aphorism: 'You shall know a word by the company it keeps'. As will be described below, we are especially interested in collocations with a grammatical dimension.

8.1. Integration of contextual information: verbs

The contextually-based approach may be exemplified by an entry from the Cambridge dictionary for the verb σῶζω, *save*, which is given in Figure 5. Because this entry is much longer than the example entries from LSJ shown in Figures 2 and 3, only its general layout is described here.

¹⁶ However, because the entries are composed in XML rather than in word-processing documents, we can insert precise line references into the XML 'master'. We plan for these to be accessible in our online edition. See FRASER (2005).

σώζω vñ. | fut. σώσω | aor. ἔσωσα | pf. σέσωκα || MID.: fut. σώσονται | aor. ἔσωσάντων || PASS.: fut. σωθήσονται | aor. ἐσώθη | pf. σέσωμαι, later σέσωμαι || This vb. is found only once in Hom. For the epic forms see σῶω. | 1 (of persons, gods, other agents) *save* (someone, sthg.) from imminent danger or death; **save, rescue** —persons, *someone's life* Hdt. S. E. + —a country, a city Hdt. S. E. + —a land (W.GEN. fr. *enemies*) S.; (transf., of a war) —a country Hdt.; (of a strategy) —a situation Th.; (of a person, a truce) —W.ACC. + DNF. (sts. W. μῆ) *someone, fr. being killed* E. || MID. **save** —one's life E. —one's eyelids (i.e. one's skin) Ar. || PASS. (of persons, cities, countries) **be saved, be safe, survive** Thgn. Hdt. Trag. +
2 **save** (someone) from an unwelcome situation; **save, rescue** —someone Trag. —W.ACC. + GEN. OF PREP. PHR. *someone, fr. troubles* S. || PASS. **be saved** —W.GEN. fr. *troubles* E.
3 **get** (sthg.) back safely; **rescue, recover** —corpses E. —(fig.) *a lost opportunity* D.
4 (of persons, gods) **bring** (W.ACC. someone) **safely or in safety** —W.ADV. OF PREP. PHR. fr. or to a place S. E. Pl. || PASS. **get safely** —fr. or to a place Hdt. Trag. + —W.DAT. to someone Theoc.
5 (of persons, gods) **preserve** (someone, sthg.) from harm; **keep safe, protect, look after** —possessions and sim. Hdt. S. E. + —persons, cities and sim. Hdt. Trag. Pl.; (of a city) —its inhabitants S. Th.; (of a class of people) —a city E.; (of a countryman) —a source of fire Od.; (of a dragon) **guard** —the golden fleece E.; (fig., of a son) —a father's

reputation E.; (of a woman) —a marriage bed (i.e. remain faithful) E.; (of a pregnant woman) —an embryo A.; (of a person) **preserve, maintain** —one's present way of thinking A. —silence E. || MID. (of a person) **keep safe** —one's treasures E. fr. —one's guard dog (fig. ref. to a person) Ar.; (of a country) —its spoils of war E.; (of a person) **preserve, maintain** —caution S. —a memory (of sthg.) E. Pl. || PASS. (of a person) **be kept safe** E.; (of inactivity, ref. to persons) **be safe** Th.

6 (of things) serve to preserve (someone, sthg.) from harm; (of obedience, caution, training, good order) **keep safe, protect** —people S. Ar. Th. X.; (of laws and customs) —a constitution, democracy Arist.; (of corks) **keep secure** —a fishing line or net A.; (of an anchor) —a ship E. fr.; (of a quiver) —an arrow E.; (of weapons) —parts of the body E.; (of time) **keep alive** —someone's hopes E.; (of the ability to light a fire) —a person S. || PASS. (of objects) **be kept safe** S. E. —W.DAT. for someone Ar.

7 **keep** to oneself; **keep safe, guard** —information, secrets A. S.

8 **preserve** (sthg.) by obedience; **observe** —instructions Trag. —laws S. E. —proverbial advice A. fr.; (hence) **fulfil** —what is ordained E.

9 **preserve** in one's mind; **remember** —someone's fortunes E.; (sts. mid.) —what one has learned E. Pl.

10 (of God, Christ) **keep safe** from eternal death; **save** —a person NT. || PASS. **be saved, be in a state of salvation** NT. —σωστέον neut. impers. vbl. adj. **it is necessary to keep safe** —weapons, people E. Ar.

This word has a wide range of meanings, comparable to the range of English *save*, and these appear in a great variety of contexts. Ten major sense sections have been identified here, and given bold translations, which are, in most instances (in sections 1, 2, 5, 6, 7, 8, 9, 10), preceded by plain-text definitions, which give more precise identification of meanings by providing a paraphrase of each sense.

In general, the order of sections follows the 'logical' ordering discussed above in 1.2, with the literal sense (*save, rescue*) in sections 1-4, a transition to stative uses (*protect, preserve*) in 5 and 6, more abstract senses (*save or remember information or instructions*) in 7-9, and finally a religious sense in 10.

The sense-sections are organised not only by the definitions and translations, but also by the verbal subjects, which are given in parentheses preceding the translations. In this entry, the subjects are usually persons or gods, though section 6 has inanimate subjects, given in general terms at the start of the section, as 'of things', and later specified in the body of that section as abstract entities (*obedience, caution, training, good order, laws, customs*) and as physical objects (*corks, an anchor, weapons*).

Although most sections have personal subjects, these are often extended to include more abstract groups, as in section 1, where the subjects include *war, a strategy, a truce*.

Following the bold translations, dependent constructions are given in italic. This gives a distinctive format to the sequence '(subject) **verb** -object', and, because the exit language, English, has the same ordering (S-V-O), it is always clear whether 'of persons' (for example) refers to a subject or to a complement. This contrasts with LSJ, whose lack of clarity in this regard is described by Glare (1987: 12-13). The regular triple ordering is preserved throughout the entry.

It may be noted that the sections are also organised internally by meaning, rather than by syntax. Middle and passive usages are given within the sections to which they belong semantically: here, in sections 1, 2, 4, 5, 6, 10. For clarity, they are given after the actives, preceded by double vertical bars.¹⁷

There is also a more general advantage to the contextual presentation: by introducing sections with verbal subjects, rather than only with the senses of the lemma, the logical principles of sense-ordering can also be applied to the grammatical context. For example, senses relating to persons and to gods are grouped together, and precede senses relating to events and logical relationships, within each numbered section, and also between sections. It may also be seen that semantic groupings dominate the structure, and that chronological information is considerably less prominent, being given only by the author abbreviations.¹⁸

8.2. Integration of contextual information: adjectives

Differences in the application of collocational information to each part-of-speech may be illustrated with an adjectival entry, for the word μέτριος, *moderate*, which is shown in Figure 6.

μέτριος ὁ οὐ (also οὐ μέτριος) *adj.* (freq. as neut.sbj. and neut.adv.; freq. in litotes) **1** of normal and not excessive size, quantity or number; (of a measure) **standard** Hdt.; (of a person) of **ordinary height** Hdt.; (of an animal) **average-sized** Plb.; (of hair) of **medium length and thickness** X.; (of a place) **medium-sized** Arist. Plb.; (of a number of people) **moderate** Pl. Arist.
2 of average and proper length (in time); (of a life-span, of a stage in life) **normal, average** S. E.fr. Pl.; (of time allotted for a speech) **appropriate** Isoc. Pl.
3 of sufficient number or quantity; (of a number of men or amount of time, needed for a task) **reasonable, appropriate** Pl. X. Plb.; (of a quantity of poison, to bring on death; of wine, not to cause intoxication) Pl. || NEG.PHR. (of the life of the unfortunate) too long S.
4 (of a person, in character, emotions, conduct) **moderate, restrained, reasonable** Thgn. Ar. Att.orats. +; (of demeanour) **unobtrusive** Pl.
5 (of a person, in way of life) without ostentation or extravagance, **moderate, not excessive** Th. E. Att.orats. Pl. +; (of dress, of quantities of food and wine, of

ceremonious occasions) Th. Pl. X. +
6 (of a deity, person or state, in the exercise of power) **moderate, fair, balanced** Th. E. Att.orats. Pl. +; (of laws, of political systems) Th. Pl. Arist.
7 (of a prayer or request) **reasonable, moderate** A. E. Att.orats. +; (of terms of an agreement) Th. D.; (of statements, discussions) Att.orats. Pl. +
8 (of means, property) **moderate, not excessive** E. Att.orats. +; (of a sum of money, of a gift) Att.orats. Pl. +; (of necessities, provisions, supplies) Pl. +
9 (of a wind, of a season, of certain times of the day) **temperate, mild** E. Ar. Pl. X.; (of an emotion or passion) **not excessive, tempered** E. D.
10 harsh but not excessively; (of a storm) **moderate** E.; (of a task, a burden, troubles) **moderate, tolerable** E. Pl. Men. Plb.; (of punishment, imprisonment, servitude) Th. Att.orats.; (of compulsion or force) Pl.
11 (pejor.) of indifferent quality or importance; (of figs) **second-rate** Hippon.; (of things said and done) **commonplace** Plb.
12 (of literary or narrative style) **moderate, balanced** Arist.

This word has a wide range of meanings (*normal, moderate, appropriate, fair, indifferent*), which vary with context. Thirteen major sense sections are given, with senses given in bold translations, which are, in half the sections (1, 2, 3, 5, 10, 11, 13), preceded by plain-text definitions.

Again, the logical ordering of senses is followed, with physical measurements preceding (spatial in 1, temporal in 2), number in 3, then references to personal character and way of life in 4 and 5, the expression of authority in

¹⁷ This format is inspired by the practice of the GI.

¹⁸ Grouping the authorial citations alongside the semantic information gives a link between chronology and meaning, creating what ZGUSTA (1989: 190, 199, 220) describes as *double articulation*.

6. In sections 7-12, we move from persons to abstract referents: requests in 7, possessions in 8, weather and emotions in 9 and 10. Sections 11 and 12 cover value-judgements, referring to objects (pejoratively) and to literary-style.

It can be seen that this sequence depends in large part on the nouns which the adjective qualifies. Consequently, the internal structure of the sections is also organised in this way, with the nouns given in parentheses preceding the translations (rather like the verbal subjects discussed in 8.1). The nouns are regularly placed at the start of each numbered section, where they delimit the semantic field for a particular sense. Alternatively, semantically-wider sections may be introduced by plain-text definitions. In section 1, for example, we can see that the sense of *standard, average* is applied to a great range of nouns, indicating humans, animals, locations, and measurements, and so a plain-text definition precedes the qualified nouns. On the other hand, the meanings in section 4 and 5 (*moderate, restrained*) can be understood only in terms of their application (to persons, in their character or way of life), and so the nouns are given first. Similar considerations apply to the *requests* in section 7 and the *possessions* in 8. In section 9, the meteorological phenomena lead to a transferred emotional sense (*temperate, tempered*).

In sum, though we follow the same logical ordering principles as developed by Liddell and Scott and by Murray, our sense-sections differ considerably in content and order from those in LSJ, because we use more detailed contextual information to identify the semantic groupings.

8.3. *Integration of contextual information: other parts of speech*

There is no space here to give further examples, but it may be seen that contextual information has a major grammatical component, and consequently that it will differ according to the part of speech of the lemma. For this reason, the Cambridge dictionary uses different entry-structures for nouns, adjectives, verbs, prepositions and adverbs. The editors can format the entries in a very regular way, specific to each part-of-speech, because we compose them in XML files (similar to the HTML used for coding web pages), which have been designed with structures matching each entry type, so giving us more precision than is possible with word-processing documents. See Fraser (2005).

9. Conclusion: lexicography, semantics and context

The contextual method was inspired by the founder of the Cambridge dictionary, John Chadwick. From his work on the *Oxford Latin Dictionary* and on the *LSJ Revised Supplement*, and also from his wartime experience as a cryptographer, Chadwick (1996: 3-6, 20-23) came to the conclusion that contextual analysis could be combined with analysis of word-form in order to arrive at the meaning of each lemma. He decided that an experimental approach could be best realised in an intermediate-sized lexicon, and since

1998, under the editorship of Dr Thompson, the Cambridge dictionary project has followed the method.

Attention to context has considerable advantages for the writers and also for readers. It gives writers an opportunity to consider new data, and so to rethink many of the semantic groupings of LSJ. Every day, we identify new senses which have not been described in the older dictionary.

Presenting meanings in their contexts is also helpful for the reader, because, as Quine (1992: 58) observed, dictionaries do not deal only in words, but also 'teach the use of sentences.' This is especially appropriate for a student dictionary, where we wish to alert readers to the subtleties of Greek word usage in the most natural way: that is, by giving examples. We have also found that giving contextual information can assist students to gain an appreciation of the ancient language, and of the ways in which its words interacted with each other in sentences, producing narrative and lyrical literature which still communicates powerfully, across the millennia, to readers today.

Bibliography

A. Dictionaries

- DGE 1980-. ADRADOS, F. R., GANGUTIA, E. *et al.*, *Diccionario griego-español*, Madrid: Consejo Superior de Investigaciones Científicas.
- GI 1995, 2004. MONTANARI, F. *et al.*, *GI: Vocabolario della lingua greca*, Turin: Loescher.
- HASE, K. *et al.* 1831-65. *Thesaurus Graecae Linguae*, Third Edition, Paris: Didot.
- LAMPE, G. W. H. (1961-8). *A Patristic Greek lexicon*, Oxford: Clarendon.
- LFgrE 1955-. SNELL, B., METTE, H. J., and MEIER-BRÜGGER, M., *Lexikon des frühgriechischen Epos*, Göttingen: Vandenhoeck & Ruprecht.
- LIDDELL, H. G. and SCOTT, R. 1843. *A Greek-English Lexicon: Based on the German work of Francis Passow*, Oxford: University Press.
- LSJ 1925-1940. LIDDELL, H. G. and SCOTT, R., revised by H. S. JONES, with the assistance of R. MCKENZIE, *A Greek-English Lexicon: a New Edition*, published in ten Parts, Oxford: Clarendon.
- LSJ Supplement 1968. BARBER, E. A., with the assistance of MAAS, P., SCHELLER, M. and WEST, M. L., *Greek-English Lexicon: a Supplement*, Oxford: Clarendon.
- LSJ Revised Supplement 1996. GLARE, P. G. W., with the assistance of THOMPSON, A. A., *Greek-English Lexicon: a Revised Supplement*, Oxford: Clarendon.
- OED 1888-1928. MURRAY, J. A. H. *et al.*, *A New English Dictionary on Historical Principles* (reprinted 1933 with the title *Oxford English Dictionary*), Oxford: Clarendon.
- OLD 1968-82. WYLLIE, J. M., GLARE, P. G. W. *et al.*, *Oxford Latin Dictionary*, Oxford: Clarendon.

- PASSOW, F. 1831. *Handwörterbuch der griechischen Sprache*, Leipzig: Vogel.
- SCHNEIDER, J. G. 1797-8. *Kritisches griechisch-deutsches Handwörterbuch: bey dem Lesen der griechischen profanen Scribenten zu gebrauchen*, Leipzig: Frommann.
- STEPHANUS 1572. ESTIENNE, H. *Thesaurus Graecae Linguae*, Geneva.
- TRAPP, E. *et al.* 1994-9. *Lexikon zur byzantinischen Gräzität*, Vienna: Österreichische Akademie der Wissenschaften.
- VALPY, A. J. and BARKER, E. H. 1816-28. *Thesaurus Graecae Linguae*, Second Edition, London: Valpy.

B. Other literature

- ABEILLÉ, A. (ed.) 2003. *Treebanks: Building and Using Parsed Corpora*, Dordrecht: Kluwer.
- ADRADOS, F. R. 1977. "Organización de los artículos del diccionario. Criterios a seguir" in F. R. ADRADOS, E. GANGUTIA *et al.* (eds.), 259-280.
- ADRADOS, F. R. 1986. "The Greek-Spanish Dictionary and Lexicographic Science", *Lexicographica* Series Maior, 2: 8-32, republished as "El DGE y la ciencia lexicográfica" in F. R. ADRADOS and J. R. SOMOLINOS (eds.), 2005, 235-244.
- ADRADOS, F. R. 1997. "More on the *Diccionario griego-español*" in H. H. HOCK (ed.), *Historical, Indo-European and Lexicographic Studies: a Festschrift for L. Zgusta on the Occasion of His 70th Birthday*, Berlin: Mouton de Gruyter, 221-231.
- ADRADOS, F. R. 2000. "La Semántica en el *Diccionario griego-español*" in M. MARTINEZ *et al.* (eds.), *Cien años de investigación semántica*, Volume 1, Madrid, 99-110, republished in F. R. ADRADOS and J. R. SOMOLINOS (eds.), 2005, 289-300.
- ADRADOS, F. R., GANGUTIA, E. *et al.* (eds.) 1977. *Introducción a la lexicografía griega*, Madrid: Institutio Antonio Nebrija.
- ADRADOS, F. R. and SOMOLINOS, J. R. 1994. "The TLG Data Bank, the DGE and Greek Lexicography." *Emerita* 62: 241-251, republished as "El banco de datos del TLG, el DGE y la lexicografía griega" in F. R. ADRADOS and J. R. SOMOLINOS (eds.), 2005, 193-203.
- ADRADOS, F. R. and SOMOLINOS, J. R. (eds.) 2005. *La lexicografía griega y el Diccionario griego-español*, Madrid: Consejo Superior de Investigaciones Científicas.
- BAMMAN, D. and CRANE, G. 2007. "Computational Linguistics and Classical Lexicography", *Digital Humanities Quarterly*, Autumn, 2007 (www.digitalhumanities.org/dhq).
- BERGK, T. 1882. *Poetae Lyrici Graeci*, Volume 3, Leipzig: Teubner.
- BERKOWITZ, L. and SQUITIER, K. A. 1990. *Thesaurus Linguae Graecae: Canon of Greek Authors and Works*, New York: Oxford University Press.
- BUCK, C. D. and PETERSEN, W. 1944. *A Reverse Index of Greek Nouns and Adjectives*, Chicago: University of Illinois Press, republished 1970, Darmstadt: Olms.
- CHADWICK, J. 1994. "The Case for Replacing Liddell and Scott", *Bulletin of the Institute of Classical Studies*, 39: 1-11.
- CHADWICK, J. 1996. *Lexicographica Graeca: Contributions to the Lexicography of Ancient Greek*, Oxford: Clarendon.
- CRANE, G. 1991. "Generating and Parsing Classical Greek", *Literary and Linguistic Computing*, 6.4: 243-245.

- DIELS, H. 1905. "Der lateinische, griechische und deutsche Thesaurus", *Neue Jahrbücher für das klassische Altertum Geschichte und deutsche Literatur*, 15: 689-697.
- FIRTH, J. R. 1957. "A Synopsis of Linguistic Theory 1930-55", *Studies in Linguistic Analysis*, Oxford: Philological Society, 1-31, reprinted in F. R. PALMER (ed.), 1968, *Selected Papers of J.R. Firth 1952-59*, London: Longmans, 168-205.
- FRASER, B. L. 2005. "Author-tagging: a Dictionary DTD as Part of the Writing Environment", *Cultural Heritage Language Technologies* (www.chlt.org/lexicon/papers/Author-Tagging.pdf).
- FRASER, B. L. 2008 forthcoming. "Beyond Definition: Organising Semantic Information in Bilingual Dictionaries", To be published in *International Journal of Lexicography*, 21; draft online at www.chlt.org/lexicon/papers/Beyond-Definition.pdf.
- GLARE, P. G. W. 1987. "Liddell & Scott: its Background and Present State" in R. BURCHFIELD (ed.), *Studies in Lexicography*, Oxford: Clarendon, 1-18.
- GLARE, P. G. W. 1997. "Liddell-Scott-Jones: Then and Now", *Hyperboreus*, 3: 205-217.
- HIORTH, F. 1955. 'Arrangement of Meanings in Lexicography.' *Lingua*, 4: 413-424.
- KARLSSON, F. et al. (eds.) 1995. *Constraint Grammar: a Language-independent System for Parsing Unrestricted Text*, Berlin: Mouton de Gruyter.
- KIPFER, B. A. 1984. "Methods of Ordering Senses within Entries" in R. R. K. HARTMANN (ed.), *LEXeter '83 Proceedings: Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983*, Tübingen: Niemeyer, 101-108.
- LOBEL, E. and PAGE, D. 1955. *Poetarum Lesbiorum Fragmenta*, Oxford: Clarendon.
- MURRAY, J. A. H. 1879. *An Appeal to the English-Speaking and English-Reading Public, to Read Books and Make Extracts for the Philological Society's New English Dictionary*, 2nd edition, Oxford: Clarendon (dictionary.oed.com/archive/appeal-1879-06/p5.html).
- MURRAY, J. A. H. 1884. "The President's Address for 1884", *Transactions of the Philological Society*, 1882-4, 501-531.
- MURRAY, J. A. H. 1888. 'General Explanations' in *OED*, Volume 1, xxiv-xxxiv.
- PANTELIA, M. 2000. "Νοῦς into Chaos: The Creation of the Thesaurus of the Greek Language", *International Journal of Lexicography*, 13.1, 1-11.
- QUINE, W. V. O. (1992) 'Meaning' in W. V. O. QUINE, *Pursuit of Truth*, Revised Edition, Cambridge MA: Harvard University Press, 37-59.
- RYDBERG-COX, J. A. 2005. "Word Profile Tool", *Cultural Heritage Language Technologies* (www.chlt.org/CHLT/Slip_Sample/index.html).
- SILVA, P. 2000. "Time and Meaning: Sense and Definition in the *OED*" in L. MUGGLESTONE (ed.), *Lexicography and the OED: Pioneers in the Untrodden Forest*, Oxford: Oxford University Press, 77-95.
- SOMOLINOS, J. R. and BERENGUER, J. A. 2005. "El trabajo de documentación en el *Diccionario griego-español*" in F. R. ADRADOS and J. R. SOMOLINOS (eds.), 2005, 105-130.
- ZGUSTA, L. 1987. "Derivation and Chronology: Greek Dictionaries and the Oxford English Dictionary (Dvādaśakośyam)", *Lexicographica Series Maior*, 23: 259-281.
- ZGUSTA, L. 1989. "The Oxford English Dictionary and Other Dictionaries (Aikakośyam)", *International Journal of Lexicography*, 2.3: 188-230.
- ZGUSTA, L. 2006. *Lexicography Then and Now: Selected Essays*, edited by F. S. F. DOLEZAL and T. B. I. CREAMER, *Lexicographica, Series Maior*, 129, Tübingen: Niemeyer.