

What the concordancer does is make the invisible visible. Patterns that would never be immediately recognisable spring to the eye with a freshness that can be quite astonishing the first time you use such a tool — and which does not lose its fascination even after long familiarity.

HARDWARE AND SOFTWARE

Hardware

If you do decide that you want to begin to use concordancing as part of a teaching or research strategy you are going to have to provide yourself with the necessary machinery and software. While you can write simple concordancers in BASIC and run them on Commodore 64's or BBC machines, you are probably going to want something like an IBM compatible PC with at least 512K of RAM (preferably the full 640K). A hard disk will make your life a lot easier, but is not essential for some software. Because of the price of Apple machines in the UK, I have no experience of working with one, but they clearly also have the potential for text analysis. You will also need access to a printer, preferably one that can produce clear condensed print (so that you can print lines longer than 80 characters).

Software

The software required for concordancing can range from do-it-yourself routines written in a programming language or as word-processing macros, to full-blown dedicated packages. In the case of the former, I have found that the programming language Snobol4 + and macros in the word-processors WordPerfect 5.1 and Microsoft Word 5 provide considerable scope for concordancing (see Tribble & Jones, 1990); in addition, Joseph Rézeau (Rézeau, 1988) has reported very interesting ways of exploiting Ashton Tate's Framework. Examples of fully developed concordancing and text retrieval systems are Electronic Text Corporation's WordCruncher and Oxford University Press's Oxford Concordancing Program. These latter programs are powerful and extremely sophisticated, but they require a hard disk and quite a long learning period before they can be used to their full extent — and they can be very frustrating for beginners.

The only concordancer of which I am aware that is specifically designed for educational use is Longman's Longman Mini-Concordancer. This has the great advantage of being very easy to use, is incredibly fast in operation and is capable of running on a modern minimum-configuration PC compatible; i.e., one floppy disk drive and 512/640K of RAM. The limitation of being able to handle up to a maximum of around 50,000 words at any one time is less significant in educational contexts than it would be in, for example, lexicographic research.

DATA

With your computer and software in place you now need data to work on. Concordancers are usually designed to work on "plain text files" (sometimes known as ASCII files). These contain no word-processing format codes and are the standard medium for computerised information interchange.

One solution to the problem of obtaining data is to buy it. Several publishing companies, including those mentioned above, are now producing collections of machine-readable text which constitute an enormously valuable resource. Alternatively, if you are in an organisation where students produce their written work on word-processors, you have a ready source of text that can be used for error analysis. There are other solutions. At a recent GAL meeting in Göttingen Joseph Rézeau described a procedure for down-loading tele-text data from BBC news broadcasts distributed through the French minitel network; a colleague of mine (Rob Ellis in Bell Language Institute, London) has just completed keying in a corpus of 40,000 words of spoken English based on transcriptions of recordings he himself has made.

Students can begin to learn directly from both their own and their fellow students' mistakes.

General and specialised corpora

The texts that you collect can be organised into general or specialised corpora. A general corpus is most useful if you wish to build up an idea of the global behaviour of vocabulary items or phrases. Specialised corpora can be particularly effective in helping students gain an idea of how language is used in a particular register of English. For example, a collection of business letters can give an invaluable insight into the types of wording and grammar appropriate to a variety of commercial transactions. A collection of short stories will, by contrast, allow for comparative studies of writers or the analysis of image and metaphor employed by particular writers or periods.

Learner text

Collecting the work of students is something that teachers have always done. Having that material in concordance (is this a new word?) form means that students can begin to learn directly from both their own and their fellow students' mistakes. Tense, order of adverbs, the behaviour of transitive verbs, article usage and so forth are all amenable to concordance study, and if students become involved in the process of formulating the questions necessary to searching through the corpus data, their awareness of language can be further enhanced.

CATEGORIES OF INFORMATION

You can look for a wide range of linguistic features with a concordancer. The set of examples below is by no means exhaustive, but it does indicate some of the potential offered by concordancing. Each example also demonstrates some of the more surprising aspects of language use that concordancing can point up. A concordance on full-stops, for example, can give insights into the thematic structure of a text or set of texts. A concordance on a single noun, e.g., "people,"

will show a range of strong adjectival collocations. Concordancing on “*ly” will show you all the derived adverbs in a text and also give learners invaluable information about word order of adverbs in natural language. In this way, the software provides students with a means for discovering or devising rules for themselves.

punctuation: [.]

...e money, and this we have now proved. "In the first year we concentrated on ...ned the boss. But no tears came yet. In the past, in the first months and eve ...ng, and admiring the other costumes. In truth, I think that my forte was a co ...aff to work in intensive-care units. Increasingly, Lord Trafford disagre ...mired, especially by old Woodfield. It gave him a feeling of deep, solid sat

single word: [people]

...kiss back pain. Four out of five people will suffer from severe and inc ...ries set up the project to give people from all walks of life the oppo ...vate party. Quite impressively, people behave as if it were one, an the ...nd us, we threaten life itself. People are waking up to this threat, a ...example, talking about the jobs people do is often used to practise de

word classes: [*ly]

...tograph on p. 108). This is most definitely an adventure lover's book, imag ...broad ink pot, and was feebly but desperately trying to clamber out again. B ...t. I'm lucky it's worked out quite differently for me. (Julie, aged eighteen ...would have done a few things quite differently, but they certainly did not co ...Work enforcing authority, so it is doubly important that you do enter any acc

verb forms: [c?me(s)/coming]

...n!" groaned the boss. But no tears came yet. In the past, in the first months ...nd as the fly tried its wings down came a great heavy blot. What would it mak ...plan outlined by the Government to come into operation from next April, could ...he charity Age Concern and it will come into force next April. Many OAPs ...ge. For younger mothers it usually comes as a surprise, an unintended consequ ...plan acc-ordinated costume, some coming as pieces of a cake which can link u

any word/part-word + specific affix: [un*ed]

...the boy except as lying unchanged, unblemished in his uniform, asleep for eve ...thought of the boy except as lying unchanged, unblemished in his uniform, asl ...unlockable poolside cubicles (most undignified) and heavily chlorinated atmos ...st few pages of this book. (For the uninitiated, riding his bicycle with a 52 ...it usually comes as a surprise, an unintended consequence of a past or curren ...e took a key off his watch-chain, unlocked a cupboard below his desk, and dr ...have been treated by qualified and unqualified people patronising and presump ...any and all words which are as yet unrecorded. With the help of Longman Wordwa

phrase of "n" words: [there is]

a, as well as to Britain. However, there is evidence of a slight increase in ...the chapel in its original state. There is a full photographic and written re ...ulic exercise equipment and sauna, there is the added appeal of the swimming ...ght is supported in the water and there is no stress on any muscle group." ...ssion at Addenbrooke's Hospital if there is enough in-terest. Cambridge

word/phrase + word/phrase separated by an indeterminate number of words: [there (are/is/were/be*)]

...lationship, and an event for which there has been little planning or preparat ...can also occur in situations where there seems to be absolutely no risk to per ...Earlier it was revealed that there had been no heart or heart-lung tran ...t fortnight. Prof Sir Roy said there needed to be twice the number of int ...ugust. The beginning of September if there is enough demand. A small charge w ...uncil refused to put up the rents, there would be no housing benefit for thou

APPLICATIONS - TWO BRIEF EXAMPLES

Example a:

The usefulness of concordancing can best be understood by looking at practical applications. At present I am working with a group of advanced students in Queen Mary and

Westfield College/Bell Language Institute's EAP programme. They continue, in spite of their general sophistication in English, to make mistakes in the use of core prepositions. This particular lesson was prepared as a paper-based exercise designed to help these learners come to grips with some of the meanings and relationships the prepositions “at” and “in” can hold. Their first task involved a conventional selective cloze activity, part of which is given in Figure 4.

Figure 4:

"For Stalin, confronted _____ August 1945 with the evidence _____ Hiroshima and Nagasaki, the outstanding fact was that the USSR possessed no strategic air force and could not deliver a direct attack _____ the United States. The best that Stalin could do was to pose a threat _____ western Europe which might deter the Americans _____ attacking the USSR."

The students were instructed to select prepositions from a limited set and attempt to “fill the gaps.” They were also instructed to mark the text with arrows which indicated whether textual information in front of, or after, the gap, determined their choice of preposition. Using such a system the first gap might produce the answer “confronted in → August 1945” but the context after “deliver” would produce “could not deliver a direct attack ← on the United States” and that after “pose” the answer “pose a threat ← to western Europe.” Discussion of the acceptability or otherwise of the students’ various choices brought home to them the dangers inherent in attempting to find translation equivalents that might serve as gap fillers. The knowledge they had to have in order to make satisfactory choices related as frequently to the collocational requirements of verbs that preceded the preposition as it did to words that followed the preposition.

This part of the class completed, students were then set to work in groups to complete exercises such as the one which is partially reproduced in Figure 5 below.

Figure 5:

In the following exercise a pair of commonly confused prepositions has been removed from the concordance lines. What prepositions are they?

| | |
|----------|----------|
| a. _____ | b. _____ |
|----------|----------|

Work in a group of two or three and decide which prepositions should be in the gaps. Are there any generalisations that you can begin to make about the way in which these two prepositions are used?

[50 examples in the original]

1. 5.4 and 5 which had first been introduced _____ 1959 and 1961 respectively. Each 55.4 an
2. ter after all. The Security Council met _____ once at the request of the United States
3. blewishes chief among them its racialism _____ home and the crassness of its admiration
4. already engineered an anti-Chinese revolt _____ the Li valley in Xinjiang (Sinkiang), w
5. ron warhead and the long range BI bomber. _____ 1979 a SALT II treaty was signed. It w
6. eeling than any decade since the fifties. _____ that earlier period, still under the sha
7. , and at a communist conference in Warsaw _____ February the Chinese were observers only
8. s became aware of it, although it was not _____ first clear to them that the Russians we
9. versing a basic tenet of American policy. _____ the Geneva conference of 1954, Zhou had
10. sed to bring the war with Japan to an end _____ a particular way. It was also a politica
11. n the cold war. Its origins are described _____ Part Six of this book. In April 1961 Ken
12. as the communes. These began to be formed _____ the beginning of 1958, were announced an
13. ss. It contained significant limitations. _____ its core was the restriction of nucle
14. csing, in December 1963, a nuclear freeze _____ Europe, which, however, the Americans di
15. bly to establish frontiers and to ensure, _____ their western end, communications with t
16. tern Meisse flows northward into the Oder _____ a point where the upstream line of the o
17. t with four of his colleagues for a month _____ the end of 1949, and were then rendered
18. ministers in 1947, at Moscow in March and _____ London in November, failed to elaborate
19. incipal leaders in the countryside, often _____ odds with his superiors in Shanghai on p
20. previous year). In negotiations in 1970 _____ Vienna and Helsinki, which occupied five

The students quickly realise that the prepositions are "in" and "at" and then, using the same "before or after" technique, they work through the exercise. Once they have made their decisions as pairs they compare their work with another group, debate differences and then check their work against a key. The final stage of work gets them to categorise the different reasons for the selection of particular prepositions. Here they will discover that some surprising contexts, for example, "at Vienna and Helsinki," are accounted for by the following term being a metonymy standing for "at the Vienna conference:" an institution rather than a physical location containing the event. Other categories are the more common "at the beginning of 1958," with at + point of time reference; "in 1959.." with in + period of time reference; and the very common "fixed phrase prepositions" such as "at odds" and "at home." Once they have completed this paper-based study, they are then grouped to investigate the use of prepositions they themselves find difficult.

In this stage they are encouraged to make contrastive studies—e.g., "of / for", "to / at" —wherever possible, recording their findings and sharing information with other groups. The availability of subject specific corpora and a quick, "in-memory" concordancer can be very helpful at this stage as students begin to come to grips with details of the rhetoric of their own area of specialisation. This type of investigative work, moving from guided exploration of the language into more free-ranging research, is an ideal example of the way a concordancer can be used in classroom language study.

Example b:

The same group in Queen Mary and Westfield College has also investigated the use of articles in English. Here the issue has not been to decide whether definite or indefinite articles should be used. Rather, it has been apparent from the group's writing that the students' problem lies in deciding whether or not there should be any article with a noun. The procedure we used in studying this problem made use of the word-list facilities of the Longman Mini-Concordancer and search and replace and macro features in WordPerfect 5.

Stage 1

The first stage in the work was a conventional consciousness-raising activity in which students had to contend with a gapped text, part of which is shown in Figure 6, where articles might or might not be needed. The discussion that arose from this exercise led to the establishing of some provisional rules about where English needs an article that could be tested during the rest of the lesson. These rules included categories such as countability and particularity, but also covered (importantly) syntactic features such as noun phrase post-modification and the plural.

Figure 6:

ARTICLES

_____ Russians decided to make _____ major issue of these developments and to resort to _____ force to stop them. They cut _____ road, rail and water routes by which _____ western occupiers communicated with _____ Berlin and stopped _____ food, electricity, gas and other necessities from being supplied regularly to _____ western sectors from _____ east. _____ legal right to use _____ routes uninterruptedly was vague - and also irrelevant in what was clearly _____ trial of _____ strength.

Stage 2

The second stage of work drew on lists of nouns derived from texts in the areas in which students had an interest. The example given below in Figure 7 was taken (using the concordancer) from an economics text book. Here students had to make decisions about the relative "countability" of the nouns in question. They were encouraged to recognise the inherent fuzziness of this category, but also to recognise that some of the nouns are central members of the set while others are peripheral. They were also encouraged to recognise the way in which abstract nouns in English are generally used uncountably. Large numbers of nouns were discussed— upwards of fifty by each group.

Figure 7:

COUNT / UNCOUNT

- 1. amount _____
- 2. analysis _____
- 3. circle _____
- 4. curve _____
- 5. demand _____

Stage 3

From here, students went on to consider the effect of post-modification on article use, this time using examples derived by concordancing on sentences. Students here had to describe the type of postmodification that had been used, labelling the sentences P [prepositional], ing ["ing" participle], ed ["ed" participle] or R [relative clause]. A sample exercise with post-modification is shown in Figure 8:

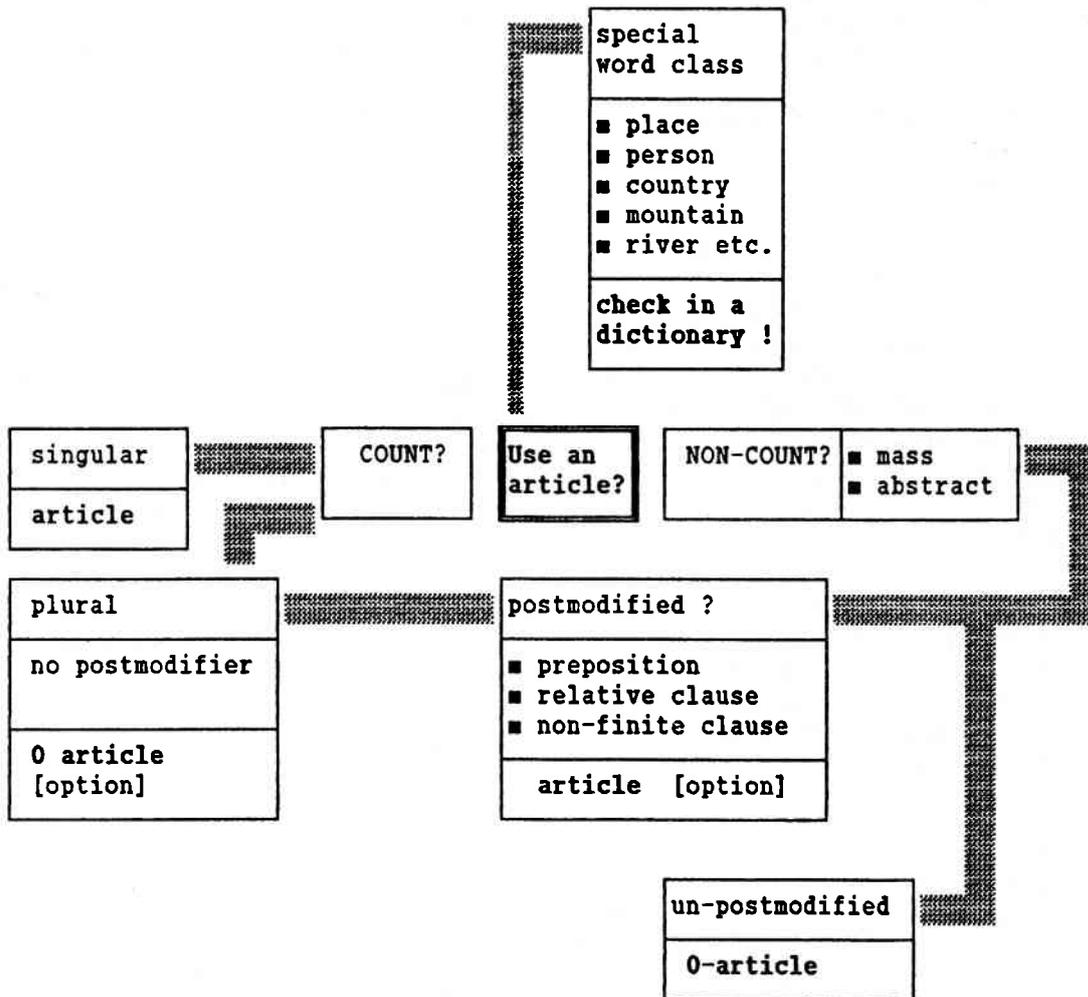
Figure 8:

POST-MODIFICATION

- 1. Between 1910 and 1914, 1930 and 1931, and 1977 and 1979 (the period of the 'Lib - Lab Pact'), the parties in office lacked a majority in the House of Commons.
- 2. These advantages range from conversation and books in the home to vacation travel and a different sense of oneself, and apply, for the children enjoying them, both in and out of school.
- 3. Whether one thinks that the differences between the programmes are big or small, or too big or too small, depends on value judgements and political perspectives.

Students were now in a position to address some of the major issues involved in the decision whether—or not—to use an article with a particular noun. The chart below presents a summary of the findings they found most helpful:

Figure 9:
ARTICLES - A POSSIBLE NETWORK OF CHOICES



Stage 4

This chart shown in Figure 10 below was then used on a concordance of nouns which had been post-edited to insert a gap at a possible article position for each line. Students had to decide what criteria applied for each case: whether it should be "article" or "0 article," and why.

Stage 5

The ideas that had been discussed and worked with during the class were then exploited in a revision activity in which authentic examples of student writing containing mistakes in the deployment of articles had to be edited by groups of two or three. Again, there was cross-comparison between the work of the groups and a great deal of discussion of the reasons for inclusion or exclusion.

Figure 10:

1. duction. It also accounts for _____ difficulty of capital theory. Becau
2. on capital and the profits of _____ enterprise (or entrepreneurship). Some
3. ce new products and so on. If _____ businessman refuses to insure against
4. between wages and the size of _____ population. Perhaps the relationshi
5. sily calculated when there is _____ monopoly as it could be when there was
6. s perfect competition between _____ buyers in the product market. Now that
7. lear because the consumer and _____ investor are the same person. Perhaps
8. ions about the firm, but hire _____ manager to run it. It is not always ea
9. demand for income in terms of _____ leisure. As in Fig. 52.3b, income is n
10. will differ in the same way. _____ factory near a large market will pay a
11. the risk of losing his money. _____ risk may be small or large, and the la
12. ion, at any given moment, and _____ people who invest at that moment are r
13. f abstinence. The notion that _____ capitalist practised virtuous thrift o
14. and indestructible powers of _____ soil is more difficult to judge. Admi
15. tter land. Obviously, land of _____ differential fertility might or might
16. l type of labour. The duty of _____ manager is to coordinate hired factors
17. ditions to, or reductions in, _____ population. It is not certain what sha
18. terest over a period of time. _____ businessman has to consider the produc
19. of the population. Perhaps _____ relationship differs between countries
20. d partly from the pressure of _____ increased population, had raised food
21. eans using lengthy processes. _____ investor will 'wait' for the machine t
22. insurance company knows that, say, 1 p _____
23. bor-rowed it has to come from _____ people whose 'rate of time preference'
24. onate increase in the size of _____ population. This is not the accepte
25. explanation. Some idea of _____ difficulty in finding an answer to thi
26. ndlord and tenant differ from _____ country to country. What we are intere
27. expensive in terms of effort (_____ leisure becomes 'dearer') and more and
28. ', he will turn to playing in _____ group if he is offered less than £40.
29. rs are unlikely to combine in _____ trade unions except with the aim of ra

This had taken a two-hour session during the first month of a 20-week course where the students attended class for two hours per week. Follow-up has continued throughout the course. Where students have continued to have problems it has been possible to discuss what is going with a shared meta-language, and now learners—especially some Thai and Punjabi speakers in the group—are producing work that is almost devoid of a fault that had crippled their performance at the beginning of the programme. There has been a great deal of work other than grammar awareness-raising in the course, this including a sustained (but implicit) process syllabus and a great deal of work on discourse and text conventions appropriate to the genres of writing required of the students. Concordancing has, however, been an invaluable tool - without which large sections of the writing course would have been impossible to achieve.

OTHER APPLICATIONS

The language-teaching potential of the combination of PC-based concordancing and electronic corpora can be seen in two lights. It will help the students by providing them with an opportunity to discover how language is used in both general and particular settings. It will also help teachers by making it possible to give authentic, intensive examples of the language features that they wish to incorporate into their syllabus. Using modern word-processors in conjunction with a concordancer, it becomes possible to prepare materials for class use that are compelling and authoritative. Using an "in-memory" concordancer and appropriate corpus material, it becomes possible for learners to find out about areas of language use that they find problematic. Given below are some of the areas in which I have started to use this tool. I will be very interested to hear what others are able to achieve as they begin to exploit this still very new approach to the very old task of teaching and learning a language.

Vocabulary

- Reinforcing new core syllabus vocabulary with lower level students
- Investigating collocation of nouns and adjectives.
- Investigating vocabulary use in specialised texts (e.g., business, medicine) both in terms of frequency and also in terms of the metaphors that are used in such genres.

Grammar

- Investigating the use of articles and determiners.
- Investigating gerunds and present participles.
- Investigating noun pre and post-modification.
- Investigating conjunctions.
- Investigating modal verbs.
- Investigating the use of sentence conjuncts.

Literature

- Character studies in prose and drama.
- Investigating recurrent images in fiction.
- Investigating rhyme patterns in poetry.
- Investigating narrative development in short stories.

CONCLUSION

As more and more sources of machine-readable text become available and as computer use in education increases, it may well be that applications programs will become the principal tools of those of us who are interested in using computers in helping our students learn languages. Word processors and data-base management software have already had a major impact on secondary and tertiary education. It is certain that text-processing software—grammar checkers, parsers and concordancers—will also have a major part to play in language teaching. Because it is relatively simple to use, does not require authoring before it can be deployed, and has such an exceptionally wide range of applications, the concordancer will perhaps be the pre-eminent software tool in this next stage in the development of computer assisted language learning. The corpora that we will need to develop in order to exploit this potential will have major significance for all of us who are interested in this application of computers in English language teaching — and perhaps for the whole English language teaching community. ☺

For more information, contact: Chris Tribble, Queen Mary and Westfield College, University of London/Bell Language Institute, Mile End Road, London, EC1 4NS, UK.

REFERENCES

Articles and books:

- Johns, T. 1986. Micro-Concord, a Language Learner's Research Tool. *System*, 14:2, 151-162.
- Johns, T. 1988. Whence and whither classroom concordancing? In T. Bongarts, et al. *Computer applications in language learning*. Foris.
- Rézeau, J. 1988. De l'utilisation d'un progiciel professionnel en E.A.O. des langues: ou comment FRAMEWORK est devenu FRAMETEACH. *Le Français dans le Monde*. August/September.
- Sinclair, J.M. 1987. *Looking up: An account of the CO-BUILD project in lexical computing*. Collins: London.
- Tribble, C.T. 1988. The use of text structuring vocabulary in native and nonnative speaker writing. *MUESLI News* (Micro Users in ESL Institutions). June.
- Tribble, C.T. and Jones, G. 1990. *Concordances in the classroom*. London: Longman (March 1990).
- Software:**
- Framework. Ashton-Tate Corporation, Maidenhead, UK.
- Longman Mini-Concordancer. Longman Group Ltd, Harlow, Essex, CM20 2JE, UK.
- Micro OCP. Oxford University Computing Service, Oxford University Press, Walton St., Oxford, OX2 6DP, UK.
- Microsoft WORD 5. Microsoft Corporation, 16011, NE 36th Way, Box 97017, Redmond, Washington, 98073-9717, USA.
- WordCruncher. Electronic Text Corp. 56000, North University Avenue, Provo, Utah, UT 84604, USA.
- WordPerfect 5.1. WordPerfect Corporation, 288 West Center Street, Orem, Utah 84057, USA.