

Pedagogical uses of monolingual and parallel concordances

Ana Frankenberg-Garcia

This paper discusses the use of concordances in the classroom, with particular reference to the pedagogical implications of the differences between parallel and monolingual concordances. Examples are given of using the two kinds of concordances in activities that involve language production, reception, correction, and testing. It is concluded that monolingual and parallel concordances have non-conflicting, complementary roles to play.

Introduction

A concordance is a list of the occurrences of a given word, part of a word, or combinations of words, together with their contexts, within a corpus of text. Figure 1, which contains concordances from the Web for the search string *a concordance is*, illustrates what a concordance is.

A concordance is a list of all the examples of a specified lexical item which occur in a corpus.

A concordance is a matchless tool for investigating texts.

Since words express ideas, themes and motifs, a concordance is highly useful in detecting patterns of meaning as well.

The student must be able to find how a word is used—and for this a concordance is invaluable.

FIGURE 1
Selected concordance lines from WebCorp¹ for *a concordance is*

Just by looking at the above concordance lines, readers can get the gist of what concordances are. Their potential for language learning is enormous, and there is already a significant body of literature demonstrating how language learners can benefit from their use (for example, Tribble and Jones 1997 and Aston 2001). Apart from being employed in the compilation of corpus-based dictionaries, grammars and syllabuses, concordances can also be utilized directly in the classroom, in an approach to language learning which, since the pioneering work that appeared in Johns and King (1991), has come to be known as data-driven learning.

This paper contrasts the use of monolingual and parallel concordances in data-driven language learning.

Monolingual concordances

Monolingual concordances such as the ones in Figure 1 can be obtained not only from the Web, where provenance is not always clear, but also from carefully designed linguistic corpora like the British National Corpus (BNC) and COBUILD's Bank of English.²

It is imperative that language learners and teachers know exactly where their concordances come from in order to know what to expect from them. English concordances from the Web, for example, may not be in standard English: anyone can publish practically anything on the Web. Concordances from carefully edited linguistic corpora are safer to use, but may in turn contain language which is too difficult for beginners or too general for advanced learners researching specialized technical topics. Users must learn to choose concordances from corpora that represent the type of language they are after, and even then it may be necessary to filter out part of the results returned.

Parallel concordances

Parallel concordances are concordances from corpora which contain texts in one language aligned with their translations into one or more languages. Parallel concordances are therefore bi- or multilingual. Figure 2 illustrates what a parallel concordance is by showing a selection of English-Portuguese concordance lines for the English negative prefix *un* taken from the Compara parallel corpus of English and Portuguese.³

| | |
|--|---|
| ...trying to get some grievance redressed, or some unjust law altered ... | ... a procurar reparar uma ofensa ou modificar uma lei injusta ... |
| ... I am an unsuitable person to be a wife. | ... sou pouco adequada para ser esposa. |
| ... he's uncommonly well set up. | ... é invulgarmente bem parecido. |
| ... he was unknown to four-fifths of Rouen | ... quatro quintos de Rouen o desconheciam |
| ... | ... |

FIGURE 2
Selected parallel concordances from Compara for *un**

Although parallel concordances are normally associated with translation studies, translator training, the development of bilingual lexicography and machine translation, several studies have referred to their potential uses in second language learning (for example, Roussel 1991 and Barlow 2000). As discussed in Frankenberg-Garcia (forthcoming), however, when using parallel concordances in the classroom it is important to bear in mind that:

- Parallel concordances are based on translations and encourage learners to compare languages. They can therefore only be appropriate in the classroom when learners share the same native language, and when it is helpful for them to use their first language as a tool for learning the second language.
- Parallel concordances provide access to so many comparable facts of linguistic performance that it is easy to lose sight of the kind of information that really matters. There is no point in sieving through parallel corpora to make second language learners focus on language

differences that do not affect their learning. Swamping learners with unsolicited language contrasts could even be detrimental to learning.⁴

- Unlike monolingual concordances, which present learners with texts written in a single language, parallel concordances contain not only two languages, L1 and L2, but also two types of language: source texts (ST) and translations (TT). When using parallel concordances for pedagogical purposes it is necessary to understand the implications of the differences between L1–L2, L2–L1, ST–TT, and TT–ST concordances.

In addition to these words of caution, it is also important to note that parallel concordances come from corpora that are representative only of (a part of) language that has been translated. Language learners and teachers cannot expect to find in them words or expressions that are beyond the scope of the texts these comparatively small corpora contain.

Having drawn attention to the need for special care when using parallel concordances for pedagogical purposes, in the next section my aim is to discuss how parallel and monolingual concordances serve different functions and have non-conflicting, complementary roles to play in the classroom. While some language learning situations call for the use of monolingual concordances, others can be dealt with more effectively by means of parallel concordances.

In either case, it is not necessary to have a computer in the classroom. Learners can use concordances by themselves, on a self-access basis, and teachers can simply bring concordance printouts to the classroom or, with the help of a word processor, turn concordance outputs into handouts.

Concordances in language production

When learners are engaged in language production activities, a question that often arises is:

How do you say _____ in L2?

The question means learners know how to say something in their L1, and are looking for a way of expressing it in L2. Monolingual concordances wouldn't be very useful here, for learners wouldn't know where to begin looking. However, used as a complement to (or instead of) bilingual and language production dictionaries, parallel concordances can help learners to find foreign words they don't know and learn in which contexts they are appropriate. Provided the parallel corpus used contains sufficient examples of the term the student is researching, an automatic search for the term in the student's L1 renders different ways of expressing it in the L2. The fact that parallel concordances offer not just linguistic equivalents, but also the contexts in which different terms are equivalent can help learners decide which term is appropriate.

Portuguese learners of English who look up *feira* in a Portuguese-English bilingual dictionary, for example, will encounter several possible translations for the word (*party, festival, feast, holiday, celebration, etc.*) and may find it difficult to choose which term to use. If they look up *feira* in the L1-L2 direction of a Portuguese-English parallel corpus like

Compara, they will be able to see not only different ways in which *feira* has been rendered in English, but also the different contexts in which each term was used (Figure 3).

| | |
|---|--|
| –Está a dar uma feira ?–perguntou. ... faz tempo que não vejo Karl Kroop nas feiras da faculdade . | ‘Having a party ?’ he said. ... you don’t often see Karl Kroop at faculty social gatherings . |
| ... semelhante ao som de um antigo sistema de altifalantes numa feira de aldeia ... | ... like the sound of an old-fashioned tannoy system at an English village fete ... |
| ... os ciclos anuais eram marcados por feiras de família ... | ... annual cycles were punctuated by family occasions ... |
| Na feira do Corpo de Deus, dizia ela, toda a aldeia da família se associava ... | At the feast of Corpus Christi, she said, her family’s whole town combined ... |
| ... os cultos, as feiras , as religiões que floresciam na sua mocidade. | ... the cults, festivals and religions that had flowered in his youth. |
| Toda a vida ela sonhara a feira . | She had dreamed of a reception all her life. |
| ... <i>Tu Bisvat</i> é o nome de uma feira judaica ... | ... <i>Tu Bisvat</i> being the name of a Jewish holiday ... |
| Ela transou com o garoto Ritchie na feira de Ano Novo. | She had it off with young Ritchie at the New Year’s Eve do . |
| ... não se queria meter em feira alheia ... | ... not wishing to interfere in other people’s celebrations ... |

FIGURE 3
Selected parallel concordances from Compara for *feira*

Another way in which parallel concordances can be useful in language production activities is in helping learners come to terms with the fact that there are certain words in their mother tongue for which there are no simple, direct translations available. Parallel concordances that go from source-texts in L1 to translations in L2 can be especially helpful when learners have to deal with culturally-bound concepts, difficult to express in the target language, like the Portuguese word *saudade* in English (Figure 4).

Another question that frequently arises when writing or speaking in a second language is:

Is it okay to say _____ in L2?

The most common way of obtaining answers to questions like these is to ask a native speaker. But native speakers may not always be available or may or may not have clear answers. Monolingual concordances from large, linguistically edited corpora like the BNC or COBUILD’s Bank of English can be very helpful in these circumstances, for they allow learners to access the combined intuitions of literally thousands of native speakers together. In contrast, parallel corpora are often not big or representative enough to provide conclusive evidence in this kind of situation.

| | |
|--|--|
| ... vamos sentir saudades de você ... | ... we shall all miss you ... |
| ... e ele sente espasmos de saudade ... | ... he has a sudden spasm of homesickness ... |
| ... uma ponta de saudade ... | ... a stab of longing ... |
| Tenho saudades tuas. | I think about you. |
| ... a lembrança da pátria que trouxe a saudade ao coração pressago. | ... the memory of my native land that brought a longing to my anxious soul. |
| Nenhuma água de Juventa igualaria ali a simples saudade . | No water from Iuventus could match simple nostalgia in that. |
| ... que saudade deixava? | ... what would he leave behind to regret ? |
| ... saudade do antigamente. | A yearning for days gone by ... |
| ... vinha-me como uma saudade dos meus tempos ocupados da repartição. | ... I felt almost nostalgic for the days when I was busy at the office. |

FIGURE 4
Selected parallel concordances from Compara for *saudade*

A student translating a Portuguese menu into English, for example, knew you roasted meat but baked bread, but was not sure whether she should say *roast* or *baked fish*. The student's searches for *roast*, *baked* and *fish* in an English-Portuguese parallel corpus (compara) did not provide her with an answer. The hits for *fish* only told her it could be cooked by frying and boiling, and the hits for *roast* and *baked* were not used in relation to *fish*. Looking up *fish* in the 100 million-word BNC Simple Search was not helpful either, for most of the 50 concordances retrieved were about live fish. However, her search for *roast* and *baked* enabled her to find information on foods that were typically baked and foods that were typically roasted, and this enabled her to conclude that *baked fish* sounded better than **roast fish* (Figures 5 and 6).⁵

If you wish to **roast** your own coffee, you should ...
The nuts are picked and dried in the sun, and then placed in a fire to **roast**.
The 32 guests enjoyed **roast** turkey ...
... to be taken with **roast** lamb...
... generous portions of **roast** chicken ...
Swathes of **roast** beef and perfect Yorkshire puddings.
... pulling **roast** chestnuts out of a grate.

FIGURE 5
Selected concordances from the BNC for *roast*

... the **baked** bread tends to collapse on cooling ...
ASADILLO (**Baked** peppers and tomatoes)
... serve sausage and **baked** beans by the fire.
A **baked** potato and some porridge, you know.
... wonderful whole **baked** fish dishes.
... a freshly **baked** baguette for breakfast.
... artichoke hearts **baked** with olive oil and garlic ...
... plum tarts and custards, crumbles and **baked** apples.

FIGURE 6
Selected concordances from the BNC for *baked*

When engaged in language production activities, parallel concordances can help learners express in L2 what they already know how to say in L1, and monolingual concordances can provide them (and non-native teachers) with answers to endless questions about language usage which require the corroboration of a native-speaker's intuitions.

Concordances in language reception

Concordances can also help learners better understand the L2. When reading in a foreign language, L2–L1 parallel concordances can help learners to understand foreign words, meanings and grammar that they are unfamiliar with. Extracting concentrated examples of parts of the foreign language that they don't quite understand, matched to equivalent forms in their mother tongue can help boost language comprehension.

For example, a Portuguese learner of English familiar with the meaning of contrast of the word *however* was having difficulty understanding its 'no matter how' meaning in the sentence *No programme, however good, can replace the role of the teacher*. He was instructed to look up *however* in a monolingual English dictionary, a bilingual English-Portuguese dictionary, an English-Portuguese parallel corpus (Compara), and a large, monolingual reference corpus of English (the BNC Simple Search). He was then asked how well each of these resources enabled him to understand that particular meaning of *however* and reported that:

- The monolingual dictionary (COBUILD, which happens to be corpus-based) helped him understand the term, but it didn't help him see what it meant in his native Portuguese.
- The bilingual dictionary (Michaelis) explained the meaning of *however* sought 'in a more direct way', but the translation given, *por mais que*, 'didn't quite fit in the sentence'.
- The monolingual concordances for *however* retrieved from the BNC were 'too many and it took too long to read them all' (Figure 7).⁶
- The parallel concordances from Compara were very helpful because they showed him 'different translations for *however* in different contexts. There were too many concordances to read, but a quick glance at the first few was enough' (Figure 8).⁷

However, this is not very easy to achieve.

Claud **however** found it increasingly hard to earn because of his Communist past and it was left to Patricia to keep the various wolves from the door.

The controlling agents of the status quo may know the power of lies; dissident sub-cultures, **however**, are closer to knowing their value'; (cited in Bronski, Culture-Clash, 41).

However, Newton's theory of light and colours, first published in 1672 and extended in his *Opticks* (1704), provided a new direction for colour theorists.

He concluded that the parallel corpus was the most useful resource, but the bilingual dictionary could save time if it contained some concrete examples. The Portuguese translations of the term seemed more important to the learner than explanations and examples in English only. For learners with similar learning styles, parallel concordances may assist language comprehension more than monolingual concordances.

FIGURE 7
First four monolingual concordances from the BNC Simple Search for *however*

FIGURE 8
First four parallel
concordances from
Compara for *however*

| | |
|---|--|
| ... there's something special about every single episode of a sitcom, however trite and formulaic it may be ... | ... há qualquer coisa de especial em cada episódio de uma série televisiva, mesmo que seja muito banal e siga sempre a mesma fórmula ... |
| Actors, however , regard writers with respect, even a certain awe. | Mas os actores têm respeito, ou até mesmo um certo temor, em relação aos autores. |
| But if it doesn't occasionally touch on the deeper, darker side of life, however glancingly, then the audience won't believe in the characters ... | Mas, se não tocar ocasionalmente no lado mais profundo e mais obscuro da vida, por muito superficialmente que seja , o público deixa de acreditar nas personagens... |
| Recently, however , it's been converted into a glazed and tiled atrium ... | Contudo , foi recentemente transformado num átrio com telhado de vidro e chão de mosaicos... |

Concordances in language correction

Concordances can also be used to help learners correct their errors. Monolingual concordances can be particularly useful in situations when learners question their teachers, requesting proof of why something was marked wrong. A Portuguese learner of English wondered why the word *informatics* in the sentence **I don't like informatics very much*, was marked wrong (to say the same in Portuguese, the learner would have used the word *informática*, which means both computers and informatics). A search for *informatics* (Figure 9), and one for *computers* (Figure 10) in the BNC helped the student establish very quickly that *computers* was more appropriate than *informatics* in the context of her sentence.

Looking up *informatics* and *computers* in a parallel corpus like Compara wouldn't be as helpful. The corpus is much smaller and contains only fiction, which is not a genre likely to contain many hits for *informatics*. Consequently, it does not provide enough examples for highlighting the difference between the two terms.

Still, parallel corpora can be very effective for dealing with L2 errors that can be traced back to the influence of the L1. Portuguese learners of English, for example, often assume that false cognates like *resume* and *resumir* mean the same. Portuguese to English and English to Portuguese searches in Compara can show these learners that, in an unmistakable way, the Portuguese *resumir* (summarize) is not the same as the English *resume* (continue) (Figure 11).

The Faculty of **Informatics** has four departments: Applied Computing, Computing Science, Information Systems, and Mathematics.

All the Faculty's students will, by the time they graduate, have been trained to bring to their work an awareness of **informatics** as a key force in social change and economic advance.

The Inter-governmental Bureau of **Informatics** was set up by the United Nations.

The systematic study of information technology (**informatics**) is in its infancy...

FIGURE 9
Selected concordances
from the BNC for
informatics

FIGURE 10
Selected concordances
from the BNC for
computers

For me, my Mont Blanc is a reminder that **computers** can't do everything and an incentive to improve my handwriting.

Hospitals now need expensive body-scanners; schools need expensive **computers**.

Computers will never replace executives, Robert Lucky reassured them.

These days everyone uses **computers**, and I didn't want to look an idiot if I told people I didn't have a clue how to work one!

FIGURE 11
Selected parallel
concordances from
Compara for *resum** (En-
Pt and Pt-En searches)

| | |
|--|---|
| He resumes work ... | Bob Busby continua o seu trabalho ... |
| He resumes his interrupted journey to the toilet ... | Conclui a viagem até ao lavabo ... |
| He turned resolutely back towards the hill and resumed his progress upwards. | Resolutamente, deu meia volta, ficou virado para a colina e retomou o caminho íngreme. |
| ... life in the plaza was about to resume . | ... a vida daquela praça estava prestes a recomeçar de novo. |
| A televisão resume-se a linhas. | That's what TV is—all lines. |
| ... para resumir , Morris Zapp não conseguia imaginar nada que quisesse e ainda não tivesse alcançado ... | ... in short , Morris Zapp could think of nothing he wanted to achieve that he hadn't achieved already ... |
| Assim lhe chamo, porque resume o universo, e o universo é o homem. | That's what I call it, because it sums up the universe and the universe is man. |
| Vou aqui resumindo , como posso, as esperanças de Natalidade. | I summarize here, as best I can, Natalidade's hopes. |

An L2 teacher whose students share the same native language is bound to find that there are many other problems of cross-linguistic influence that parallel concordances can help sort out.

Concordances in language exercises and testing

It is often the case that when language teachers are preparing exercises or tests for their students, they need to find or invent sentences which focus on specific words or grammar structures. While invented sentences often sound artificial and stilted, finding authentic sentences without the help of a corpus can be a daunting task. Both parallel and monolingual concordances make it very easy for the teacher to locate such structures in the context of authentic sentences, which they can then use to prepare tests or exercises for their students. Figure 12 contains a mix of monolingual concordances from the BNC containing adverbs of frequency, which can be easily deleted and used as a word-order exercise that focuses on the position of those adverbs in the sentence. Figure 13, in turn, contains parallel concordances from Compara for the Portuguese preposition *com*. The English part of

FIGURE 12
Selected concordances
from the BNC for *always*,
never, and *often*

Leading me where I have **always** wanted to go yet never known how.
But **often** people stayed at home, finding jobs to do around the house.
The reader is **never** bored—; but could be disappointed that the book does not go beyond its 120 pages.
Little Billy **always** told lies to his mother, but they were **never** convincing.

FIGURE 13
Selected parallel
concordances from
Compara for *com*

| | |
|--|--|
| ... milhares de pessoas morriam diariamente com SIDA em África ... | ... thousands died daily of AIDS in Africa ... |
| ... não podia contar com o apoio dele. | ... I couldn't count on any real support from him. |
| estou preocupada com o Joe. | 'I'm worried about Joe.' |
| —Adeus—respondeu Alistair com uma voz de alívio. | 'Bye,' Alistair said in a voice loud with relief. |

the concordance can help Portuguese learners of English to see that the Portuguese preposition *com* is not always equivalent to *with*, its most common English equivalent.

As many language teaching materials available today are meant for learners of different first language backgrounds, parallel concordances are particularly useful to meet a demand for language exercises geared to monolingual settings, where learners share similar L1-related difficulties.

Conclusion

This paper has attempted to show different uses of parallel and monolingual concordances in second language learning. While in some situations parallel concordances are more relevant, in others monolingual ones are more appropriate. It is believed that the two types of concordances have non-conflicting, complementary roles to play. With common sense and practice, it is not hard to establish whether or not concordances are suitable, and if so, which concordances are best for which situations.

Revised version received February 2004

Notes

- 1 WebCorp allows use of the World Wide Web as a corpus. Access is online and free at <http://www.webcorp.org.uk/>
- 2 The BNC offers free online access to its Simple Search facility at <http://sara.natcorp.ox.ac.uk/lookup.html>, where any person can access up to 50 concordance lines from a 100 million word collection of written and spoken British English from the early nineties. At a small cost, the complete service is also available online or on CD-ROM. The Bank of English is made up of 450 million words of contemporary British and American English (written and spoken). Its demo facility can be used free of charge at

<http://www.collins.co.uk/corpus/CorpusSearch.aspx>. A paid subscription service is also available.

- 3 Free online access to parallel concordances is available for the following language pairs:
English–Portuguese (+ 1.7 million words of published fiction)
<http://www.linguateca.pt/COMPARA/>
English–French, English–Spanish, English–German, and English–Danish (World Health Organization and European Union documents)
<http://khnt.hit.uib.no/webtce.htm>
English–Slovene (one million words of technical documents, fiction, and official texts by the Slovene government)
<http://nl2.ijs.si//index-bi.html>

English–Chinese (300,000 English words and 500,000 Chinese characters of mostly legal and documentary texts)

<http://www.edict.com.hk/concordance/paralleltxts/>

English-Italian (700,000 English words from the Brown corpus translated into Italian)

<http://multisemcor.itc.it/frameset2.php>

It is also possible for language teachers to build small, home-made parallel corpora using software like Multiconcord (<http://web.bham.ac.uk/johnstf/lingua.htm>) or ParaConc (<http://www.ruf.rice.edu/~barlow/parac.html>).

- 4 The Contrastive Analysis Hypothesis was criticized precisely for this; not all differences between languages are relevant to second language learning (Wardhaugh 1970).
- 5 The (corpus-based) *Oxford Collocations Dictionary* would also have answered this query satisfactorily.
- 6 Only the 34th of the 50 concordances supplied contained the ‘no matter how’ meaning of *however*. The learner did not get as far as reading it. His search would have been more successful had he tried the more sophisticated part-of-speech query: *however* + *ADJ*.
- 7 The fact that the first four concordances gave the student the answer he needed was fortuitous. The ‘no matter how’ meaning of *however* could have turned up much further down in the list of concordances retrieved.

References

- Aston, G.** (ed.). 2001. *Learning with Corpora*. Houston: Athelstan.
- Barlow, M.** 2000. ‘Parallel texts in language teaching’ in S. Botley, T. McEnery, and A. Wilson (eds.). *Multilingual Corpora in Teaching and Research*. Amsterdam: Rodopi, 106–115.
- Frankenberg-Garcia, A.** 2004. ‘Lost in parallel concordances’ in G. Aston, S. Bernardini, and D. Stewart (eds.). *Corpora and Language Learners*. Amsterdam: John Benjamins.
- Johns, T.** and **P. King** (eds.). 1991. *Classroom Concordancing*. Birmingham: The University of Birmingham Centre for English Language Studies.
- Roussel, F.** 1991. ‘Parallel concordances and tonic auxiliaries’ in T. Johns and P. King (eds.).
- Tribble, C.** and **G. Jones.** 1997. *Concordancing in the Classroom: A Resource Guide for Teachers*. Houston: Athelstan.
- Wardhaugh, R.** 1970. ‘The contrastive analysis hypothesis’. *TESOL Quarterly* 4: 123–30.

The author

Ana Frankenberg-Garcia holds a PhD in Applied Linguistics from the University of Edinburgh, and is an auxiliary professor at ISLA, in Lisbon. She is joint project leader of the COMPARA corpus, and her current research interests focus on the use of corpora for language learning and translation studies.

Email: ana.frankenberg@sapo.pt