Concordances, Collocations and Lexical-based Language Analysis: Implications for Computer-aided Pedagogy

Preston Perluss, University of Paris-IV

Abstract
Current personal-computer technology allows every computer user to perform statistical and concordance-type studies of written documents. How can teachers best benefit from such technology? What methods and hypotheses appear the most useful to develop educational materials using these techniques? Will they bring about an overall improvement in foreign language teaching? Such will be the object of consideration of this paper. The fundamental role played by a "collocation" will be illustrated through a wide variety of examples. The contextual study of words will draw our attention to their often unpredictable behavior and shed light on the weaknesses besetting "standard grammars" and their associated pedagogies.

Preamble

How often has a student asked about the meaning of a given word – let's say "welfare"? Without hesitation the teacher will reply, "it depends on the context." This contextual dependence raises pedagogical questions to which modern technology — in the even modest form of personal computers — might be of some avail. Personal computers offer the potential for enormous change in both classroom teaching strategies and student apprenticeship techniques. The most appropriate use for our high-tech tools awaits discovery and development. Herein I merely seek to display the range of present-day possibilities. For my own research I have used a Macintosh Duo 280c but kindred software exists for systems running DOS Windows. The lexical analysis examples presented here have all been produced by the software application Hyperbase© available from Professor Étienne Brunet, à l’UFR Lettres 98 bd Herriot 06204 Nice. A version for DOS equally exists.

Several hypotheses concerning language learning and language teaching have recently come to the fore: first, the study of disembodied sentences, and a fortiori words, torn from context, provides little means of advancing a learner's use (and not mere knowledge) of language which is inherently context dependent.¹ Thus, the need arises for a coherent way of reconstructing context while at the same time analyzing some particular linguistic phenomenon, so to discern regularities in usage.

Moreover, the very idea of a word might well merit reconsideration when we take into account several facts concerning both speech and writing². To dodge the problem of specifying a definition for "word", researchers use the term "lexie" or "lexeme" to describe the functional unit of language "memorized in competence"³ that is, the base unit used by native speakers. Without begging any questions, it seems reasonable to accord credence to the idea that people do not systematically reconstruct all they say or write from a mass of items commonly referred to as words⁴. There exist larger blocks of language, learnt as wholes, knit together by hidebound convention. And while the productive features of language allow us to create novel

¹Support for this view arises from considerations of synonymy, for one example see [Carter and McCarthy: 1988 p. 29]. Michael Lewis pleads strongly for context-based teaching [Lewis:1993 p.80-81]
²Bally [Bally: 1951] in his classic treatise draws attention to the conceptual obscurity surrounding the definition of the term "word"; much more recently Tournier [Tournier: 1991] makes a similar observation
³"Competence" for those unfamiliar with syntactic grammars is a fundamental term used to describe a speaker's knowledge of the language as a mechanism for creating all possible sentences of that language. See [Brown, Malmkjær and Williams: 1997]

TESOL-France http://www.tesol-france.org/
utterances from our linguistic inventory, both the inventory and these same features must first be acquired through contextual familiarity. Finally, a large part of our commonplace intuitions about words derives from a skill we all take so much for granted we are no longer aware that it is acquired: I refer to literacy, our ability to read. My goal, when bringing to bear computer-aided lexical analysis, is to develop a general approach (I avoid saying method) which will allow us, as teachers, to analyze texts written for specific purposes and thus create classroom materials while simultaneously coming to grips with the lexical niceties of English, in particular collocational unity.

If faithfully presenting realistic context to our students is our primary concern, then we must draw on true-life materials. From these materials we must extract the essential lexical structures to teach the students in the given time frame which is allotted to us. We must go for the essential. Thus, we must find the essential. First, I wish to introduce a technical term "vocable" which may appear at first superfluous, but will (I hope) little by little gain plausibility: by "vocable" I mean any series of letters separated by blanks or punctuation marks in a text. Obviously, this concerns the written or transcribed language. The word "word", I insist, hides a lush undergrowth of references and I wish to bypass these for a moment.

What we need to establish within the framework of a given text or discursive domain, are the most commonly-used expressions whether we consider them to be institutionalized utterances, functional blocks (called "polywords" by Michael Lewis), or merely compounds formed from sequences of vocables.

More importantly, and more generally, we need to establish a high-fidelity sample of what is written and said. And we need to display this to the students who will observe, formulate interpretations and implement their hypotheses through application to other situations and contexts. But first and foremost, to understand some particular segment of the language, the learners must see the segment in context, at work, in a real linguistic situation. If we can exhibit a certain number of situations the learner will all the more readily generalize from each occurrence. We must forever bear in mind this fact. No one sentence can confirm and refute an interpretation of the various uses its component parts might have. But when you display a larger array of possibilities to the learners, the learners themselves are ideally placed to appreciate the nuances and possible variations to which a given vocable might be subject. We appeal to an inductive approach to language learning rather than merely a deductive approach wherein either the teacher or a dictionary provides a rule which is then exemplified and finally practiced only to be embarrassingly (for the teacher) overturned by some example further on. A final remark should also guide us: the actual language appears more chaotic and unpredictable than the simplified structures we all tendentiously teach in the classroom. Yet students must learn to structure that chaos.

Some examples

Consider a text. Most of what follows applies to limited corpora. In fact, it seems all the more reasonable to limit the scope of lexical analysis to texts or conversations chosen for some specific purpose. We must stress the importance of context and the context may well be the text itself (the expression "cotext" applies here) or a given conversation whose intentional circumstances are fully explored. First, we shall seek to extract those vocables which occur the most frequently in a text and which allow us to predict something of the overall information context of the text. The larger the sample or corpus, the greater the risk that the most frequent vocables be those with the lowest informative value precisely because they occur in such diverse contexts. Recall Lewis's law: the greater the number of contexts in which a vocable is found, the greater the need for a context to understand that vocable.

Figure 1 (below) displays a typical exercise based on the extraction of a low-information value vocable from a text. We have made a concordance of "how". By concordance is meant the simultaneous listing of all occurrences of a given vocable or series of vocables in a specified corpus.

---


TESOL-France http://www.tesol-france.org/
Consider the use of "how" in the following sentences. "How" is often given as the English equivalent for the French "comment"; but what does this vocable mean in a given context? Are "how" and "comment" merely used in questions? In what other contexts do they occur?

Consider the use of "how" in the following sentences. "How" is often given as the English equivalent for the French "comment"; but what does this vocable mean in a given context? Are "how" and "comment" merely used in questions? In what other contexts do they occur?

P2  3a| raises the question of how any society can function
P2  3c| a limited number of themes: how the mechanization of work has been
P3  5b| of machines determines how men think, act and dream. ¶ In
P3  5c| with ever greater subtlety how demand, supply and price interact
P4  8h| had not left their mark on how workers behave both on the job and
P9  18b| them to decide what work to do, how to do it and even when to do it.
P9  18d| greater autonomy in their work. How these demands will be reconciled

"How", thus, replaces the relative expression "the way in which" in the above contexts.

Consider the use of "how" in the following sentences. "How" is often given as the English equivalent for the French "comment"; but what does this vocable mean in a given context? Are "how" and "comment" merely used in questions? In what other contexts do they occur?

Figure # 2, a window from the application Hypertext®, lists the most frequently occurring vocables in a given text by order of frequency. What is the subject of the text and by what chain of thought do you reach your conclusions? The reader can make an educated guess on the basis of the primarily nominal clues provided. The lexical base of a text is composed of the most frequently occurring informative vocables. Obviously, frequency must be interpreted as being relative to the size of the text. The text under study contains approximately 8924 separate vocable occurrences (certain of these are punctuation marks) with only 1754 vocable forms of which some 900 occur only once (the latter are called, in linguistic jargon, the text’s hapax).
Figure # 2

<table>
<thead>
<tr>
<th>rang</th>
<th>frq mot</th>
<th>rang</th>
<th>frq mot</th>
<th>rang</th>
<th>frq mot</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>793</td>
<td>31</td>
<td>31</td>
<td>51</td>
<td>control</td>
</tr>
<tr>
<td>2</td>
<td>388</td>
<td>32</td>
<td>31</td>
<td>52</td>
<td>work</td>
</tr>
<tr>
<td>3</td>
<td>308</td>
<td>33</td>
<td>29</td>
<td>53</td>
<td>technology</td>
</tr>
<tr>
<td>4</td>
<td>253</td>
<td>34</td>
<td>29</td>
<td>54</td>
<td>so</td>
</tr>
<tr>
<td>5</td>
<td>265</td>
<td>35</td>
<td>29</td>
<td>55</td>
<td>other</td>
</tr>
<tr>
<td>6</td>
<td>517</td>
<td>36</td>
<td>29</td>
<td>56</td>
<td>only</td>
</tr>
<tr>
<td>7</td>
<td>200</td>
<td>37</td>
<td>29</td>
<td>57</td>
<td>most</td>
</tr>
<tr>
<td>8</td>
<td>177</td>
<td>38</td>
<td>28</td>
<td>58</td>
<td>was</td>
</tr>
<tr>
<td>9</td>
<td>153</td>
<td>39</td>
<td>28</td>
<td>59</td>
<td>resource</td>
</tr>
<tr>
<td>10</td>
<td>113</td>
<td>40</td>
<td>28</td>
<td>60</td>
<td>process</td>
</tr>
<tr>
<td>11</td>
<td>109</td>
<td>41</td>
<td>27</td>
<td>61</td>
<td>must</td>
</tr>
<tr>
<td>12</td>
<td>101</td>
<td>42</td>
<td>26</td>
<td>62</td>
<td>made</td>
</tr>
<tr>
<td>13</td>
<td>83</td>
<td>43</td>
<td>26</td>
<td>63</td>
<td>its</td>
</tr>
<tr>
<td>14</td>
<td>80</td>
<td>44</td>
<td>25</td>
<td>64</td>
<td>company</td>
</tr>
<tr>
<td>15</td>
<td>79</td>
<td>45</td>
<td>25</td>
<td>65</td>
<td>were</td>
</tr>
<tr>
<td>16</td>
<td>78</td>
<td>46</td>
<td>22</td>
<td>66</td>
<td>controlled</td>
</tr>
<tr>
<td>17</td>
<td>03</td>
<td>47</td>
<td>22</td>
<td>67</td>
<td>when</td>
</tr>
<tr>
<td>18</td>
<td>62</td>
<td>48</td>
<td>22</td>
<td>68</td>
<td>they</td>
</tr>
<tr>
<td>19</td>
<td>61</td>
<td>49</td>
<td>21</td>
<td>69</td>
<td>robots</td>
</tr>
<tr>
<td>20</td>
<td>56</td>
<td>50</td>
<td>21</td>
<td>70</td>
<td>operations</td>
</tr>
<tr>
<td>21</td>
<td>53</td>
<td>51</td>
<td>21</td>
<td>71</td>
<td>has</td>
</tr>
<tr>
<td>22</td>
<td>52</td>
<td>52</td>
<td>21</td>
<td>72</td>
<td>could</td>
</tr>
<tr>
<td>23</td>
<td>53</td>
<td>53</td>
<td>20</td>
<td>73</td>
<td>but</td>
</tr>
<tr>
<td>24</td>
<td>44</td>
<td>54</td>
<td>20</td>
<td>74</td>
<td>been</td>
</tr>
<tr>
<td>25</td>
<td>44</td>
<td>55</td>
<td>20</td>
<td>75</td>
<td>which</td>
</tr>
<tr>
<td>26</td>
<td>43</td>
<td>56</td>
<td>19</td>
<td>76</td>
<td>disc</td>
</tr>
<tr>
<td>27</td>
<td>39</td>
<td>57</td>
<td>19</td>
<td>77</td>
<td>about</td>
</tr>
<tr>
<td>28</td>
<td>37</td>
<td>58</td>
<td>19</td>
<td>78</td>
<td>up</td>
</tr>
<tr>
<td>29</td>
<td>35</td>
<td>59</td>
<td>19</td>
<td>79</td>
<td>through</td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td>60</td>
<td>19</td>
<td>80</td>
<td>three</td>
</tr>
</tbody>
</table>

91  13  processing  121  9  their  151  8  before  
92  13  inventory  122  9  program  152  8  back  
93  13  if  123  9  plant  153  8  areas  
94  13  and  124  9  percent  154  8  aircraft  
95  12  some  125  9  linked  155  7  workpiece  
96  12  small  126  9  line  156  7  wood  
97  12  photograph  127  9  however  157  2  where  

TESOL-France  http://www.tesol-france.org/
Lexical analysis provides a rapid insight into not so much the precise argumentation used in a text, but a skeletal representation of its language. The essential lies therein. The analysis helps both the student and the teacher. In a classroom situation, whereas the initial study of a text or discipline will embrace the nominal structures of the thematic language, the teacher's fundamental ambition must be to find a suitable number of verbs associated with these nominal structures. Let's say that the nominal structures form the abstract description of the discipline and the verbs put these abstractions into specific situations. The verbs, in a sense, rely on the intentional context of the speakers and listeners. When designing a course syllabus most of the language-based exercise material will be found within the text corpora furnished by the students or clients.

Figure #3 provides a good example of text-specific collocation: What nominal structures co-occur with assess? What compounds are formed? “Assessing” and “risk” co-occur as do “assessment” and “information”. The trainees will be called upon to note the compound structures such as “supervisory assessment” as well as “assessing the impact of...” In this example, readers should notice that the concordance is based on a form “assess” whose morphology varies. The verb, noun and participle are all used. Can the student and teacher apply their skills to draw general conclusions concerning the uses of the different forms?

Figure #3

1a] a framework for supervisors to assess information about the
1b] the framework provides for the assessment of supervisory information
1c] ( e . g . for the purpose of assessing market risks and amines ) .
1d] that supervisors can use in assessing the impact of derivatives on
1e] not specified , allowing for the assessment of information obtained
1f] may be obtained and assed through on - site examinations
1g] that for purposes of assessing an institution’s market risk
1h] supervisors to assess the risks of derivatives and
1i] that supervisors can use in assessing the impact of derivatives on
1j] that for purposes of assessing an institution’s market risk
2a| a framework for supervisors to assess information about the
2b| the framework provides for the assessment of supervisory information
2c| ( e . g . for the purpose of assessing market risks and amines ) .
2d| that for purposes of assessing an institution’s market risk
2e| a framework for supervisors to assess information about the
2f| the framework provides for the assessment of supervisory information
2g| ( e . g . for the purpose of assessing market risks and amines ) .
2h| that for purposes of assessing an institution’s market risk
3a| conceptual methods for assessing the risk exposures related to
3b| conceptual methods for assessing the risk exposures related to
3c| conceptual methods for assessing the risk exposures related to
3d| conceptual methods for assessing the risk exposures related to
3e| conceptual methods for assessing the risk exposures related to
3f| conceptual methods for assessing the risk exposures related to
3g| conceptual methods for assessing the risk exposures related to
3h| conceptual methods for assessing the risk exposures related to
3i| conceptual methods for assessing the risk exposures related to
3j| conceptual methods for assessing the risk exposures related to
4a| have identified as important for assessing the risks arising from firms’
4b| have identified as important for assessing the risks arising from firms’
4c| have identified as important for assessing the risks arising from firms’
4d| have identified as important for assessing the risks arising from firms’
4e| have identified as important for assessing the risks arising from firms’
4f| have identified as important for assessing the risks arising from firms’
4g| have identified as important for assessing the risks arising from firms’
4h| have identified as important for assessing the risks arising from firms’
4i| have identified as important for assessing the risks arising from firms’
4j| have identified as important for assessing the risks arising from firms’
5a| covers row approaches for assessing market risks . ¶One is to
5b| allows the supervisor to assess the concentration of credit risk
5c| allows the supervisor to assess the concentration of credit risk
5d| covers row approaches for assessing market risks . ¶One is to
5e| allows the supervisor to assess the concentration of credit risk
5f| allows the supervisor to assess the concentration of credit risk
5g| allows the supervisor to assess the concentration of credit risk
5h| allows the supervisor to assess the concentration of credit risk
5i| allows the supervisor to assess the concentration of credit risk
5j| allows the supervisor to assess the concentration of credit risk
6a| of information important for assessing the impact of derivatives on
6b| of information important for assessing the impact of derivatives on
6c| of information important for assessing the impact of derivatives on
6d| of information important for assessing the impact of derivatives on
6e| of information important for assessing the impact of derivatives on
6f| of information important for assessing the impact of derivatives on
6g| of information important for assessing the impact of derivatives on
6h| of information important for assessing the impact of derivatives on
6i| of information important for assessing the impact of derivatives on
6j| of information important for assessing the impact of derivatives on
7a| the supervision assessment of the information necessary
7b| the supervision assessment of the information necessary
7c| the supervision assessment of the information necessary
7d| the supervision assessment of the information necessary
7e| the supervision assessment of the information necessary
7f| the supervision assessment of the information necessary
7g| the supervision assessment of the information necessary
7h| the supervision assessment of the information necessary
7i| the supervision assessment of the information necessary
7j| the supervision assessment of the information necessary
8a| the supervision assessment of the information necessary
8b| the supervision assessment of the information necessary
8c| the supervision assessment of the information necessary
8d| the supervision assessment of the information necessary
8e| the supervision assessment of the information necessary
8f| the supervision assessment of the information necessary
8g| the supervision assessment of the information necessary
8h| the supervision assessment of the information necessary
8i| the supervision assessment of the information necessary
8j| the supervision assessment of the information necessary
9a| be accessible to supervisors to assess the risks of derivatives and
9b| be accessible to supervisors to assess the risks of derivatives and
9c| be accessible to supervisors to assess the risks of derivatives and
9d| be accessible to supervisors to assess the risks of derivatives and
9e| be accessible to supervisors to assess the risks of derivatives and
9f| be accessible to supervisors to assess the risks of derivatives and
9g| be accessible to supervisors to assess the risks of derivatives and
9h| be accessible to supervisors to assess the risks of derivatives and
9i| be accessible to supervisors to assess the risks of derivatives and
9j| be accessible to supervisors to assess the risks of derivatives and
10a| the framework discusses ways to assess the concentration of credit risk
10b| the framework discusses ways to assess the concentration of credit risk
10c| the framework discusses ways to assess the concentration of credit risk
10d| the framework discusses ways to assess the concentration of credit risk
10e| the framework discusses ways to assess the concentration of credit risk
10f| the framework discusses ways to assess the concentration of credit risk
10g| the framework discusses ways to assess the concentration of credit risk
10h| the framework discusses ways to assess the concentration of credit risk
10i| the framework discusses ways to assess the concentration of credit risk
10j| the framework discusses ways to assess the concentration of credit risk
11a| data on derivatives should be assessed with sufficient frequency to
11b| data on derivatives should be assessed with sufficient frequency to
11c| data on derivatives should be assessed with sufficient frequency to
11d| data on derivatives should be assessed with sufficient frequency to
11e| data on derivatives should be assessed with sufficient frequency to
11f| data on derivatives should be assessed with sufficient frequency to
11g| data on derivatives should be assessed with sufficient frequency to
11h| data on derivatives should be assessed with sufficient frequency to
11i| data on derivatives should be assessed with sufficient frequency to
11j| data on derivatives should be assessed with sufficient frequency to
12a| allow independent supervisory assessment of an institution’s market
12b| allow independent supervisory assessment of an institution’s market
12c| allow independent supervisory assessment of an institution’s market
12d| allow independent supervisory assessment of an institution’s market
12e| allow independent supervisory assessment of an institution’s market
12f| allow independent supervisory assessment of an institution’s market
12g| allow independent supervisory assessment of an institution’s market
12h| allow independent supervisory assessment of an institution’s market
12i| allow independent supervisory assessment of an institution’s market
12j| allow independent supervisory assessment of an institution’s market
13a| conceptual methods for assessing the risk exposures related to
13b| conceptual methods for assessing the risk exposures related to
13c| conceptual methods for assessing the risk exposures related to
13d| conceptual methods for assessing the risk exposures related to
13e| conceptual methods for assessing the risk exposures related to
13f| conceptual methods for assessing the risk exposures related to
13g| conceptual methods for assessing the risk exposures related to
13h| conceptual methods for assessing the risk exposures related to
13i| conceptual methods for assessing the risk exposures related to
13j| conceptual methods for assessing the risk exposures related to
14a| with the aim of achieving the assessment of understandable and
14b| with the aim of achieving the assessment of understandable and
14c| with the aim of achieving the assessment of understandable and
14d| with the aim of achieving the assessment of understandable and
14e| with the aim of achieving the assessment of understandable and
14f| with the aim of achieving the assessment of understandable and
14g| with the aim of achieving the assessment of understandable and
14h| with the aim of achieving the assessment of understandable and
14i| with the aim of achieving the assessment of understandable and
14j| with the aim of achieving the assessment of understandable and
15a| needed for supervisory assessment . ¶Each of the four broad
15b| needed for supervisory assessment . ¶Each of the four broad
15c| needed for supervisory assessment . ¶Each of the four broad
15d| needed for supervisory assessment . ¶Each of the four broad
15e| needed for supervisory assessment . ¶Each of the four broad
15f| needed for supervisory assessment . ¶Each of the four broad
15g| needed for supervisory assessment . ¶Each of the four broad
15h| needed for supervisory assessment . ¶Each of the four broad
15i| needed for supervisory assessment . ¶Each of the four broad
15j| needed for supervisory assessment . ¶Each of the four broad
16a| to an effective supervisory assessment of the credit risk inherent
16b| to an effective supervisory assessment of the credit risk inherent
16c| to an effective supervisory assessment of the credit risk inherent
16d| to an effective supervisory assessment of the credit risk inherent
16e| to an effective supervisory assessment of the credit risk inherent
16f| to an effective supervisory assessment of the credit risk inherent
16g| to an effective supervisory assessment of the credit risk inherent
16h| to an effective supervisory assessment of the credit risk inherent
16i| to an effective supervisory assessment of the credit risk inherent
16j| to an effective supervisory assessment of the credit risk inherent
17a| may also be necessary for assessing liquidity risk . ¶In addition
17b| may also be necessary for assessing liquidity risk . ¶In addition
17c| may also be necessary for assessing liquidity risk . ¶In addition
17d| may also be necessary for assessing liquidity risk . ¶In addition
17e| may also be necessary for assessing liquidity risk . ¶In addition
17f| may also be necessary for assessing liquidity risk . ¶In addition
17g| may also be necessary for assessing liquidity risk . ¶In addition
17h| may also be necessary for assessing liquidity risk . ¶In addition
17i| may also be necessary for assessing liquidity risk . ¶In addition
17j| may also be necessary for assessing liquidity risk . ¶In addition
18a| positions . Supervisors should assess information on market risk by
18b| positions . Supervisors should assess information on market risk by
18c| positions . Supervisors should assess information on market risk by
18d| positions . Supervisors should assess information on market risk by
18e| positions . Supervisors should assess information on market risk by
18f| positions . Supervisors should assess information on market risk by
18g| positions . Supervisors should assess information on market risk by
18h| positions . Supervisors should assess information on market risk by
18i| positions . Supervisors should assess information on market risk by
18j| positions . Supervisors should assess information on market risk by

TESOL-France http://www.tesol-france.org/
Here, in figure #4, we return to the text whose lexical frequency hierarchy was given in figure 1. Note the compounds in which machine occurs.

**Figure #4**

P1 1d| by the industrial robot , a machine
designed to replace the
P1 3d| , building materials and machine
tools . ¶It is in these
P2 4a| shall frequently refer to is the machine
shop . There metal parts are
P2 4a| on a lathe . ¶The same set of machine
tools can serve to make a great
P2 4b| the efficient organisation of machine
shop operations difficult .
P2 4c| or numerically controlling , the machine
tools themselves were
P3 7a| to determine how a cutting machine
such as a lathe , must be
P3 7a| the capacity of the cutting machine
the material from which the
P3 7b| the machinist set up his machine
according to drawings supplied
P3 7b| ; when numerically controlled machine
tools were introduced , the
tool to make it illustrates
P3 8a| a part and in programming a machine
were kept only by the
P3 8c| for processing on a particular machine
P3 8e| of materials , the control of machine
tools and other single -
P4 14c| of labor , materials , machine
time and other resource
P6 18b| card can be imprinted with a machine
readable code such as the
P7 20c| earliest numerically controlled machine
tools were programmed by means
P7 20c| tape . Each instruction to the machine
was represented by a pattern of
P7 20d| digital computer mounted on the machine
 ¶A modern computer -
P7 20e| the computer control enables the machine
from cutting too deep into the
P7 20f| in some cases it can signal the machine
operator to change or sharpen
P7 21a| - numerically controlled machine
tools are linked by a hierarchy
P7 21a| machine tools . Typically each machine
is controlled by a
P7 21b| mainframe computer to any of the machine
tools in the network . ¶In
P7 21c| about the status of each machine
the volume of its production
P7 21c| controllers . As many as 100 machine
tools can be connected in such
P7 21d| must still be moved from one machine
tools are further linked by a
P7 21d| direct - numerically controlled machine
spends cutting metal can be as
P7 21f| , the fraction of each shift a machine
 tool standing alone the cutting
P8 22a| products . A worker operating a machine
Here in figure 5 we have an exercise using a more grammatical vocable "one". Find the fixed expressions using "one". The idea of grammaticalized lexis gains currency through the use of such methods: it is not by merely learning the formal grammar of the vocable “one”, that a learner may acquire the vocable’s bewildering uses: disembodied analysis daunts more than it describes.

**Figure #5**

How do we use the vocable “one”? Does any expression come to mind? Consider the following concordance of “one”. What can be said about the lexical unit in which it occurs? Are there sentence frames involved or polywords?

P1 1c| turns out that manufacturing is **one of the most** difficult sectors of
P1 1d| to replace the production worker **one for one**. ¶Actually the direct work
P1 1d| the production worker **one for one**. ¶Actually the direct work of
P2 5a| as it **passes from one person to another**. ¶The time
P2 5e| and the routing of a part **from one process to another** on the factory
P3 8e| ¶By linking the six areas one can achieve what Joseph Harrington
P4 10b| - view drawing. ¶In return **one of the most** important benefits to
P4 10c| - ray - tube terminal. In one method, called finite - element
P4 11b| custom valves from six months to one month. A manufacturer of molds for
P5 15e| , for moving the part **from one operation to the next**

P6 P7 | **One of the most important benefits** of a
P6 P7 | **must still be moved from one machine to another by manual**
P8 22d| by hand, with each worker doing one small step of the job and passing
P8 24a| with other machines. ¶**One of the most important applications**
P9 25d| ¶Computers can then be linked to one another and to the central data
P9 27b| workers on the first shift, one worker on the second shift and )
P9 27b| workers on the second shift and ) one on the third. ¶Such small
P0 28a| ¶One advantage of the computer
P0 31c| bulwarks can be machined **at one time**. Each cutting head of the
P0 32a| a later stage in the assembly. One advantage of the computer
P0 32b| that carries workpieces from one machine to the next The
P0 33c| is represented schematically in one of the positions it assumes
Figure 6 provides the teacher with the raw material for various exercises — exercises which one might not have otherwise contrived, had we not performed the correspondence for “time(s)”. What verbs collocate with “time”? Which multi-vocable blocks appear useful for expressing time relations between the present and past? What compound nouns are formed from “time”? Students can be led to discern the difference between “time” used as a nominal base in compound nouns and when used as a modifier in such compounds. Finally, the learners (and teachers) can compare the plural and singular forms of “time”. How often do we teach “time” with an “s” attached? — as a verb? Or in fixed expressions such “as many times as you like”? Does this list exhaust the manifold potential of the vocable “times”?

Figure #6

P1  1g| part to spend 95 percent of the time required for its manufacture
P2  4c| themselves were introduced some time ago and have been widely adopted .
P2  5a| one person to another . ¶The time required for a memorandum to
P2  6d| it could be recalled at any time . ¶
P4  11b|. ¶Another company - reduced the time needed to design custom valves
P4  12b| process characteristics as the time required for the setup of the
P5  14a| manufacturing process at a given time . ¶For example , a manufacturing
P5  14c| of labor , materials , machine time and other resource elements that
P5  15d| assembly of a product , on the time needed for manufacturing each part
P5  15d| part ( including not only the time spent actually working on the part
P5  15e| working on the part but also the time needed for setting up machines ,
P5  15e| at each station ) , on the lead time needed for purchasing parts from
P5  17c| it must take into account the time needed for shipment . It can , be
P6  18a| . ¶For example , a worker’ s time card can be imprinted with a
P6  18b| working hours recorded by the time clock are automatically assigned
P7  19b| workers rather than spending his time in efforts to relieve shortages of
P7  20e| around several axes at the same time ; the computer control enables the
P7  21g| tool standing alone the cutting time may be as low as 10 to 30 percent
P0  29f| along the applied force is 100 times the distortion in other
P0  29f| Setup times are so reduced that such a system

TESOL-France http://www.tesol-france.org/
Figure 7 provides an example of a multiple context cloze, easily constructed with concordancing software. Such exercises, often used in the past to draw on students' memories so as to test and reinforce prior material, can be adjusted to actually teach new contexts and new applications for well-know vocables. Multiple context clozes remove the same form from diverse contexts and thus stress the range of application a given vocable or group of vocables (polyword or even an institutionalized utterance) might have. In the following example the missing vocable “order” has an overwhelming frequency in the polyword “in order to” but that frequency does not exclude other types of occurrence such as “a single order for”.

Figure #7

P2 5c| concerns had become onerous. In ___ to get work done on lime the
P2 6c| of a cathode - ray tube in ___ to " draw, " or enter geometric
P4 11d| it was originally intended. In ___ to design a new part and plan how
P5 15b| automatically generates an ___ for the wood on August 4 , four
P5 15d| uncertainties in demand. %In ___ to introduce manufacturing
P6 16c| . In the Kanban system the ___ for a part to be made at one
P6 16d| set in motion by a single ___ for . finished products at the
P7 19b| relieve shortages of parts , to ___ repairs for machines and to
P0 33d| precise geometry of the part in ___ to program the cutting tool , can
P0 33d| is needed must be done in ___ to determine when the component
P0 35e| back - scheduled deadlines in ___ to conform to the master schedule

Figure 8 furnishes a highly informative example of one of the most common ways to express logical consequence: by means of the three vocable unit (polyword) "as a result". Students will probably notice the presence of the comma after the polyword, providing further proof of the polyword’s innate independence. As well, note the cases where “of” links the polyword to a complex noun group which expresses the cause. The examples have been taken from two different texts.

Figure #8

dw 7d| agricultural sector. %As a result, prices for these products shot
dw 8b|Israel devoid of anything. As a result they probably had negative
dw 9f|rather than the usual two. %The results have not been long to
dw 11a| this situation and that, as a result, the %account deficit is
mg 17a| of 1997 and into 1998. As a result, we expect the yield curve to
mg 21d| t income during 1996 ). %As a result, we are sceptical regarding the
mg 25c| by the markets. %As a result, Steve also remains relatively
mg 26d| fully discounted, and, as a result, sees German bond yields rising
mg 31a| for some months yet, and, as a result, sees official interest rates
mg 31c| in corporate issuance. %As a result, the flow of funds is likely to
mg 31c| in corporate cash flow. %As a result, we still expect a balance of
mg 32b| in both cases, and, as a result, expect yields to rise ( in the
mg 37c| ( especially ) the banks. %As a result, it is difficult to find
mg 39c| a dramatic profit recovery as a result of its recent investment in new
mg 40b| of at least 15%, mainly as a result of the continued recovery of B&Q
mg 42a| where growth is strongest. As a result, we think it has a higher share
mg 48a| unlikely to deteriorate on the result of the US election. It was less

Figure 9, a simple listing of the occurrences of “it” within two successive paragraphs, serves as an example for students to demonstrate that “it” seldom changes reference within a single paragraph. Another interesting exercise consists in displaying all the contexts containing “it” within a text. The various roles fulfilled by the vocable “it” provide food for thought as to the nature of textual cohesion and coherence. The same, of course, might be done for all relative pronouns.

Figure #9

Because the design is simple to alter in electronic form IT can be changed as many times as necessary without the major effort of redrawing. The design is accessible to everyone who must work with IT as soon as IT is electronically filed, so that manufacturing functions such as the planning and scheduling of production can be started earlier.

TESOL-France http://www.tesol-france.org/
This method of scheduling is called distribution resource planning, and IT must take into account the time needed for shipment. IT can, be employed to generate shipment dates for various products; IT’s output can become input for a manufacturing-resource-planning system.

As a further example of the possibilities arising from concordance-based investigations of language, examine the respective contexts in figures # 10A & 10B for “those” and “these”. If someone asked you cold out what functional characteristics tell them apart, I fear that the “extra-linguistic”, distance-from-speaker-to-reference-object criterion may indeed be adduced. Reality justifies a totally different criterion (a primary criterion, however, which is not without counter examples): in text “these” serves to anaphorically tag noun groups while “those” refers to some set of persons introduced in the text. Note that although “those who” is an intensely powerful collocation; “these who” is not at all a collocation, I would go so far as to affirm its impossibility.

**figures # 10A & 10B**

P2 3a| A disquieting feature of these dynamic internal shifts in the
P3 6b| some of the more important of these misperceptions and the inadequate
P4 7g| and take proper account of these developments, the performance of
P5 10c| operating at full capacity, these delaying tactics have in certain
P6 11g|1940 the proportions employed in these kinds of occupation were
P8 15a| These last considerations are a bridge
P9 18d| autonomy in their work. How these demands will be reconciled with
P9 18g| of workers by inflation makes these employees receptive to union
P0 19d| foreign investments, excluding these fees, came to $32.8 billion,
P0 21d| at or below the poverty level. ¶These trends can be disregarded only by

P1 1f| unemployment destabilized those societies, that national policy
P3 6e| added to the personal wealth of those in power. Only in retrospect has
P5 9b| authority and legitimacy. ¶Only those economists who believe everything
P6 11f| members of the work force: those between 21 and 42 years of age
P6 12a| increase in the proportion of those in the 25-to-29 age group who
P6 12b| and the large fraction of those families that live at or below
P0 21c| and the large fraction of those families that live at or below
P0 24b| mining and construction) and those in the service sector, defined
P0 24c| is narrowly defined to designate those workers who provide services
In conclusion

The full pedagogical value of this kind of lexical analysis remains to be exploited. The method of learning a language through collocational study reproduces, in a more condensed version, our own apprenticeship with written language. Both students and teachers may avail themselves of this new technology to learn the lexical mysteries or regularities which each text harbors. Ultimately, every learner should have access to a library-like data-base which can be analyzed accordingly, that is by performing concordances upon selected parts, when the need arises. Lectiel at the École Normale Supérieur de Saint Cloud is one example of such a library. The Trésor de la Langue Française has such a data base (FRANTEXT) which may be consulted from work-stations in various university libraries; unfortunately, individual taxpayers cannot access this data base from home. In the future, students in every discipline will perform corpus analyses on the literature in a specific domain in order to learn the vocabulary thereof. More generally, the concordancing techniques will continue to evolve with the addition of statistical tools to refine search methods. Concordancing and lexical analysis offer a solution to one of the most thorny problems in language learning: the mastery of collocations, networks of associations much too vast for any single dictionary to embrace. Thus, in the near future, alongside the dictionary will sit the work-station ready and able to flesh out all too abstract definitions.

Biodata

Preston PERLUSS majored in mathematics and minored in linguistic philosophy at San Francisco State University. In 1985, he obtained a magistère in teaching French as a foreign language from the University of Paris IV-Sorbonne. Since then, he has been teaching and conducting research in lexicology and discourse analysis while preparing a doctorate at the University of Paris-IV in 18th century Parisian urban history.

6The system of computer-aided learning as presented by Mr Arnaud Pelfrène during the colloquium on Usage des Nouvelles Technologies dans l'enseignement des Langues, Université de Technologie de Compiègne, 13-14 March 1997.
Bibliography

General works on vocabulary, lexicology, and “lexical philosophies”
Bally, Charles. Traité de Stylistique français, 1951 Paris, Klincksieck
Goody, Jack. La Raison Graphique : la domestication de la pensé sauvage, 1979, Paris, les Editions de Minuit. (Describes the deep modifications the introduction of literacy exercises on our culture and cultural perception)
Nattigner, James R., & DeCarrico, Jeanette S. Lexical Phrases and Language Teaching, Oxford University Press
Perluss, Preston, "Lexical-based Language Learning", in Prism, vol. 1 Summer 1997

Works on Corpus linguistics
Barlow, Michael, homepage internet site :http://www.ruf.rice. edu/barlow/corpus.html