A Corpus Analysis of an Anatomy Textbook:
Preliminary Findings and Implications for
Medical English Materials Development

Simon FRASER
Walter DAVIES
Keiso TATSUKAWA
Institute for Foreign Language Research and Education
Hiroshima University

In this article, we describe the creation of a corpus and word lists using Gray’s Anatomy for Students. The research is part of a wider project being carried out in collaboration with the medical faculty at Hiroshima University. A major goal of the project is to develop a lexically-based syllabus for medical students based on a corpus of carefully selected medical texts (see Davies, Fraser, Lauer, & Howell 2013; Fraser, 2013). Our initial intention was to focus on the construction of a corpus of medical research articles, in the belief that learners need to be equipped with the English skills necessary to engage in and comprehend scientific research at the global level. However, feedback from the medical faculty has highlighted the necessity for learners to first gain a thorough understanding of the subjects underpinning medical knowledge (see Davies, Fraser, & Tatsukawa, 2014, elsewhere in this journal). This means that there is a need for pedagogical materials which can help students at the early stages of their studies to not only acquire the most important terms but, equally importantly, to learn how to use these words.

We wanted to investigate, therefore, whether the corpus linguistic techniques and methodology used in the analysis of research articles to create medical word lists (see, e.g., Fraser, 2013; Wang, Liang & Ge 2008) could be applied to a major medical textbook. The feedback we have received from the medical faculty has stressed the importance of anatomy, and Gray’s Anatomy for Students was chosen for corpus treatment. Our aim was two-fold: to create lists of the most useful words in anatomy based on their frequency of occurrence, and to investigate possible ways in which the corpus findings might be used in the development of classroom materials that accurately reflect medical English in use.

ANATOMY: A SUITABLE CANDIDATE FOR CORPUS ANALYSIS?

Although there have been a number of corpus-based lexical studies in the past few years investigating the most frequently used lexical items in medical disciplines (e.g., Baker, 1988; Fraser, 2007, 2009; Wang et al., 2008), these have made use of corpora of scientific articles. A recent exception is Hsu (2013), who created a list of sub-technical medical items from a textbook corpus spanning 31 medical subject areas; however her analysis did not involve a detailed
examination of a particular sub-discipline. Chung and Nation (2003, 2004) carried out a corpus analysis of an anatomy textbook, but their aim was to determine the role of technical vocabulary rather than produce lists of words for learners. To this end, they determined that anatomy does indeed contain a very high proportion of technical words (almost one-third of the total number of running words), many of which are based on Greek or Latin roots. This finding was replicated for pharmacology by Fraser (2005).

Anatomy, then, with its high technical load, would not seem to be particularly suitable for corpus treatment. As Lowe (2010: 5) puts it, when comparing anatomy with other medical sub-disciplines: “Anatomy is a special case; full of nouns, the labels for structures and parts... [it] is probably not the most helpful class for analysis.” However, this should not deter us from trying; Fraser (2013), for instance, has shown that it is possible to create a highly efficient list of core vocabulary for pharmacology, a subject which also contains a high proportion of nouns in the form of drug nomenclature.

CREATING AN ANATOMY CORPUS

The anatomy textbook selected was Gray’s Anatomy for Students, which is highly recommended as a reference book for medical students. This textbook is an “easy-to-understand”, clinically-focused version of Gray’s Anatomy, considered to be one of the most influential works on the subject. Numbering 1,136 pages, it consists of the following nine chapters: The Body; Back; Abdomen; Pelvis and Perineum; Lower limb; Upper limb; Head and Neck; Surface anatomy; Clinical cases.

Although a digital version of Gray’s Anatomy for Students is available, we decided to carry out the present analysis using the paper version. This enabled us to familiarize ourselves with the scanning and text preparation techniques which we will need in future analyses. The textbook was prepared for scanning using an electric paper cutter to remove the binding and separate the pages. Individual pages were then fed automatically into a scanner and converted into digital (PDF) format. Adobe Acrobat OCR recognition software was used to convert the PDF files into text format ready for editing and corpus analysis. Separate files were created for each chapter of the book.

ANALYZING THE CORPUS

Creating a Frequency List

AntConc 3.2.4m, a freeware corpus analysis program designed by Laurence Anthony (see Anthony, 2005, for details), was used to produce lists of the most frequent words and to create concordances to provide contextual information. The corpus consists of 361,087 words in total, and was found to contain 10,919 word types. Even for a subject where a large number of terms would not be unexpected, the word type total is suspiciously high; by way of comparison, Chung and Nation (2003) found there to be just 4,270 word types in Clinically Oriented Anatomy. This discrepancy may simply turn out to be explained by the fact that the corpus requires further “cleaning up” to remove misspelled words and other artefacts created by the digitization process.
Nevertheless, it is something that requires further investigation.

What are the Most Frequent Words?

AntConc produced a list of all the words in the corpus ordered according to frequency. The unit of counting was the individual word form, as it was felt that the use of lemmas or word families could conceal important information about how the words are used in the text. Table 1 lists the 100 most frequent words; Appendix 1 shows the most frequent 500 words. Function words (e.g., prepositions, articles, pronouns, conjunctions) were included in the lists, as their frequency of occurrence provides useful insight into the nature of anatomy.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>THE</td>
<td>46419</td>
</tr>
<tr>
<td>2</td>
<td>OF</td>
<td>18400</td>
</tr>
<tr>
<td>3</td>
<td>AND</td>
<td>13328</td>
</tr>
<tr>
<td>4</td>
<td>TO</td>
<td>6964</td>
</tr>
<tr>
<td>5</td>
<td>IS</td>
<td>5383</td>
</tr>
<tr>
<td>6</td>
<td>A</td>
<td>5112</td>
</tr>
<tr>
<td>7</td>
<td>IN</td>
<td>5030</td>
</tr>
<tr>
<td>8</td>
<td>FIG</td>
<td>3132</td>
</tr>
<tr>
<td>9</td>
<td>NERVE</td>
<td>2974</td>
</tr>
<tr>
<td>10</td>
<td>FROM</td>
<td>2792</td>
</tr>
<tr>
<td>11</td>
<td>ARTERY</td>
<td>2379</td>
</tr>
<tr>
<td>12</td>
<td>ANTERIOR</td>
<td>2325</td>
</tr>
<tr>
<td>13</td>
<td>POSTERIOR</td>
<td>2280</td>
</tr>
<tr>
<td>14</td>
<td>ARE</td>
<td>2195</td>
</tr>
<tr>
<td>15</td>
<td>WITH</td>
<td>2024</td>
</tr>
<tr>
<td>16</td>
<td>BY</td>
<td>1995</td>
</tr>
<tr>
<td>17</td>
<td>MUSCLE</td>
<td>1922</td>
</tr>
<tr>
<td>18</td>
<td>LATERAL</td>
<td>1862</td>
</tr>
<tr>
<td>19</td>
<td>WHICH</td>
<td>1753</td>
</tr>
<tr>
<td>20</td>
<td>ON</td>
<td>1722</td>
</tr>
<tr>
<td>21</td>
<td>IT</td>
<td>1720</td>
</tr>
<tr>
<td>22</td>
<td>SUPERIOR</td>
<td>1604</td>
</tr>
<tr>
<td>23</td>
<td>MUSCLES</td>
<td>1594</td>
</tr>
<tr>
<td>24</td>
<td>INFERIOR</td>
<td>1512</td>
</tr>
<tr>
<td>25</td>
<td>MEDIAL</td>
<td>1372</td>
</tr>
<tr>
<td>26</td>
<td>INTO</td>
<td>1360</td>
</tr>
<tr>
<td>27</td>
<td>THROUGH</td>
<td>1278</td>
</tr>
<tr>
<td>28</td>
<td>SURFACE</td>
<td>1231</td>
</tr>
<tr>
<td>29</td>
<td>THAT</td>
<td>1215</td>
</tr>
<tr>
<td>30</td>
<td>PART</td>
<td>1196</td>
</tr>
<tr>
<td>31</td>
<td>AS</td>
<td>1163</td>
</tr>
<tr>
<td>32</td>
<td>BETWEEN</td>
<td>1110</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rank</th>
<th>Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>THIS</td>
<td>1097</td>
</tr>
<tr>
<td>34</td>
<td>BONE</td>
<td>1037</td>
</tr>
<tr>
<td>35</td>
<td>FOR</td>
<td>1014</td>
</tr>
<tr>
<td>36</td>
<td>AT</td>
<td>1002</td>
</tr>
<tr>
<td>37</td>
<td>LEFT</td>
<td>967</td>
</tr>
<tr>
<td>38</td>
<td>WALL</td>
<td>965</td>
</tr>
<tr>
<td>39</td>
<td>RIGHT</td>
<td>961</td>
</tr>
<tr>
<td>40</td>
<td>NERVES</td>
<td>943</td>
</tr>
<tr>
<td>41</td>
<td>VEN</td>
<td>921</td>
</tr>
<tr>
<td>42</td>
<td>BRANCHES</td>
<td>920</td>
</tr>
<tr>
<td>43</td>
<td>CAVITY</td>
<td>881</td>
</tr>
<tr>
<td>44</td>
<td>DEEP</td>
<td>862</td>
</tr>
<tr>
<td>45</td>
<td>BE</td>
<td>850</td>
</tr>
<tr>
<td>46</td>
<td>OR</td>
<td>850</td>
</tr>
<tr>
<td>47</td>
<td>AN</td>
<td>841</td>
</tr>
<tr>
<td>48</td>
<td>JOINT</td>
<td>829</td>
</tr>
<tr>
<td>49</td>
<td>HEAD</td>
<td>817</td>
</tr>
<tr>
<td>50</td>
<td>SIDE</td>
<td>808</td>
</tr>
<tr>
<td>51</td>
<td>LIGAMENT</td>
<td>806</td>
</tr>
<tr>
<td>52</td>
<td>BRANCH</td>
<td>792</td>
</tr>
<tr>
<td>53</td>
<td>TWO</td>
<td>762</td>
</tr>
<tr>
<td>54</td>
<td>NECK</td>
<td>745</td>
</tr>
<tr>
<td>55</td>
<td>UPPER</td>
<td>729</td>
</tr>
<tr>
<td>56</td>
<td>PASS</td>
<td>727</td>
</tr>
<tr>
<td>57</td>
<td>LOWER</td>
<td>695</td>
</tr>
<tr>
<td>58</td>
<td>THESE</td>
<td>691</td>
</tr>
<tr>
<td>59</td>
<td>VEINS</td>
<td>673</td>
</tr>
<tr>
<td>60</td>
<td>PELVIC</td>
<td>663</td>
</tr>
<tr>
<td>61</td>
<td>MAY</td>
<td>657</td>
</tr>
<tr>
<td>62</td>
<td>ITS</td>
<td>651</td>
</tr>
<tr>
<td>63</td>
<td>EACH</td>
<td>644</td>
</tr>
<tr>
<td>64</td>
<td>ARTERIES</td>
<td>621</td>
</tr>
<tr>
<td>65</td>
<td>FOSSA</td>
<td>611</td>
</tr>
<tr>
<td>66</td>
<td>ABDOMINAL</td>
<td>604</td>
</tr>
<tr>
<td>67</td>
<td>BODY</td>
<td>595</td>
</tr>
<tr>
<td>68</td>
<td>CAN</td>
<td>594</td>
</tr>
<tr>
<td>69</td>
<td>FIBERS</td>
<td>591</td>
</tr>
<tr>
<td>70</td>
<td>REGION</td>
<td>590</td>
</tr>
<tr>
<td>71</td>
<td>VESSELS</td>
<td>568</td>
</tr>
<tr>
<td>72</td>
<td>THORACIC</td>
<td>560</td>
</tr>
<tr>
<td>73</td>
<td>ALSO</td>
<td>548</td>
</tr>
<tr>
<td>74</td>
<td>SPINAL</td>
<td>548</td>
</tr>
<tr>
<td>75</td>
<td>SMALL</td>
<td>540</td>
</tr>
<tr>
<td>76</td>
<td>INTERNAL</td>
<td>539</td>
</tr>
<tr>
<td>77</td>
<td>MAJOR</td>
<td>532</td>
</tr>
<tr>
<td>78</td>
<td>MIDDLE</td>
<td>531</td>
</tr>
<tr>
<td>79</td>
<td>FORAMEN</td>
<td>527</td>
</tr>
<tr>
<td>80</td>
<td>EXTERNAL</td>
<td>507</td>
</tr>
<tr>
<td>81</td>
<td>WAS</td>
<td>505</td>
</tr>
<tr>
<td>82</td>
<td>PROCESS</td>
<td>499</td>
</tr>
<tr>
<td>83</td>
<td>MARGIN</td>
<td>485</td>
</tr>
<tr>
<td>84</td>
<td>VERTEBRAL</td>
<td>482</td>
</tr>
<tr>
<td>85</td>
<td>PASS</td>
<td>478</td>
</tr>
<tr>
<td>86</td>
<td>SUPPLY</td>
<td>478</td>
</tr>
<tr>
<td>87</td>
<td>CANAL</td>
<td>476</td>
</tr>
<tr>
<td>88</td>
<td>FASCIA</td>
<td>473</td>
</tr>
<tr>
<td>89</td>
<td>PLEXUS</td>
<td>469</td>
</tr>
<tr>
<td>90</td>
<td>SUPERFICIAL</td>
<td>464</td>
</tr>
<tr>
<td>91</td>
<td>BLOOD</td>
<td>463</td>
</tr>
<tr>
<td>92</td>
<td>VIEW</td>
<td>455</td>
</tr>
<tr>
<td>93</td>
<td>TENDON</td>
<td>453</td>
</tr>
<tr>
<td>94</td>
<td>STRUCTURES</td>
<td>436</td>
</tr>
<tr>
<td>95</td>
<td>OTHER</td>
<td>433</td>
</tr>
<tr>
<td>96</td>
<td>THEY</td>
<td>431</td>
</tr>
</tbody>
</table>
Excluding function words, the top ten words in the anatomy textbook are *fig* (figure), nerve, anterior, posterior, muscle, lateral, superior, muscles, inferior, and medial. That *fig* is the most frequent is unsurprising, as descriptions in anatomy naturally rely heavily on labelled diagrams. The other words can all be considered to be key anatomical terms and positional words. In addition to occurring frequently, all of them are found in each chapter of the book. The fact they have wide distribution in addition to high frequency suggests that these are words at the core of the discipline.

**What Kind of Words are They?**

Perhaps the first, and most reassuring, observation that can be made about the most frequent words in the list is that they are not all the overtly technical words that we might expect. Most of the words in the top 100, with the exception of perhaps only *foramen* and *fossa*, will at least be recognizable to the layperson, and there are few highly specialized terms of Greco-Latin origin. However, we have to be careful because, as concordance patterns make clear, many of these words take on a technical meaning when combined with other words. Also, the further we proceed down the list, the more words sourced from Greek or Latin we find (see Appendix 1).

Many words, such as nerve, artery, blood, and muscle might be considered "lay-technical": terms which are obviously central to the medical field, but whose basic meaning can be understood by someone without expert knowledge. However, we also find words which are technical in the sense that they are everyday words taking on an additional, specialized meaning in a medical context. Fraser (2012) labels these polysemous words "cryptotechnical": examples from the anatomy word list include superior, inferior, process, margin, cavity, wall, and aspect. Cryptotechnical words are important because of their potential for confusion. Words like these often become "technicalized" when combined with other words: take superior articular process, for example, a highly technical term even though both superior and process can be found in any general word list. The fact that we find many cryptotechnical and lay-technical words is in accord with Chung and Nation’s (2003) and Fraser’s (2005) findings that specialized texts contain a much higher proportion of technical words than had previously been thought (35% of the total words in anatomy and 37% of the words in a pharmacology textbook were found to be technical).

In the top 100 list, the finding that 57 words are nouns would probably be expected, but there is also a surprisingly high (36) number of adjectives in the list: these include the important positional words occurring in pairs of polar opposites such as anterior-posterior, superior-inferior, lateral-medial, and internal-external. Just outside the top 100 we find proximal-distal and dorsal-ventral. We observe that each of the words in these pairs occurs with a very similar frequency to its counterpart. An interesting exception is dorsal-ventral, where dorsal occurs 281 times in the corpus, but ventral is found only 38 times, a much lower frequency.
When we investigate the top 500 words (see Appendix 1), we see the trend continuing; almost one in two of these words is a noun (45.4%), but as many as one in four (24.2%) is an adjective. The relatively high proportion of adjectives supports our understanding that the study of anatomy involves not just the listing of structures and parts, but also descriptions of the appearance and positions of these parts ("bony protuberance", "middle lobe").

Prepositions, too, would appear to have an important role to play in anatomy. Comparison with a general frequency list (created from the Brown Corpus, and available with AntConc) shows that the prepositions from, by, into, through, and between all have high "keyness" values (i.e., they occur with much higher frequencies than we would expect in a general corpus), confirming the importance of spatial descriptions in anatomy. Typical examples from the concordance are "musculature from the ileum continues into each flap" and "drainage from the gastrointestinal system passes through the liver".

Other than forms of the verb be, the only two verbs occurring in the top 100 are pass and supply. (Interestingly, it was expected that branches, occurring 920 times in the corpus, would be used at least some of the time as a verb, but it was only found as a noun.) That few verbs are found is perhaps unsurprising, since anatomy is the study of structures, positions and relationships rather than processes. However, the high frequencies of pass and supply (both occur 478 times) suggest the importance of these words, particularly as found in pass through and supply the muscle/tissue, in describing anatomical systems and the connections between them. Other "connecting verbs" found in the top 500 are connect, occur, descend, attach, cross, and relate.

Multiword Units

As mentioned previously, in anatomy words like superior, aspect, and process take on a technical sense when combined with other words. To further investigate this phenomenon, AntConc’s "N-grams" tool was used to provide a list of the most frequent multiword sequences. The majority of sequences generated were bigrams (two words), with a smaller number consisting of three words; very few sequences are longer than this. Table 2 lists the 100 most frequent two-word combinations, which include major muscle and middle ear, as well as the highly technical sciatic foramen and infratemporal aorta. (See Appendix 2 for the top 300 two-word terms, and Appendix 3 for the top 100 three-word terms.)

| 1. MARGIN OF                      | 8. ANTERIOR TO                      |
| 2. ASPECT OF                      | 9. COMPARTMENT OF                   |
| 3. ASSOCIATED WITH                | 10. INFERIOR TO                     |
| 4. REGIONAL ANATOMY              | 11. BODY OF                         |
| 5. POSTERIOR TO                  | 12. CAROTID ARTERY                  |
| 6. BORDER OF                     | 13. MAY BE                          |
| 7. ABDOMINAL WALL                | 14. CONSISTS OF                     |
15. ORIGINATES FROM VENA CAVA
16. PASSES THROUGH ATTACHED TO PELVIC CAVITY
17. POSTERIOR SURFACE PASS THROUGH LATERAL TO MEDIAL SIDE
18. ATTACHED TO ORAL CAVITY
19. PASS THROUGH VERTEBRAL COLUMN ATTACHMENT OF LOWER LIMB
20. TEMPORAL BONE THORACIC WALL
21. ANTERIOR SURFACE GLUTEAL REGION
22. MESENTERIC ARTERY UPPER LIMB
23. NASAL CAVITY ANTERIOR RAMI
24. BRANCHES FROM SUPPLY TO CRANIAL FOSSA
25. VERTEBRAL COLUMN JUGULAR VEIN LATERAL SIDE
26. ANTERIOR SURFACE SCULAR FORAMEN LATERAL SIDE
27. TEMPORAL BONE FIBULAR NERVE LATERAL SURFACE
28. THORACIC WALL RADIAL NERVE LATERAL SURFACE
29. ANTERIOR SURFACE INFRATEMPORAL FOSSA CRANIAL FOSSA
30. GLUTEAL REGION INTERNAL CAROTID LATERAL WALL
31. MESENTERIC ARTERY INFRATEMPORAL FOSSA POSTERIOR COMPARTMENT
32. UPPER LIMB RADIAL NERVE VAGUS NERVE
33. NASAL CAVITY INTERNAL JUGULAR POSTERIOR COMPARTMENT
34. REYNIL ARTERY INGUINAL RING FLEXOR DIGITORUM
35. BRANCHES FROM SUPPLY TO CRANIAL FOSSA INFERIOR VENA
36. JUGULAR VEIN INFERIOR VENA TEMPORAL BONE
37. SCIULAR FORAMEN MEDIAN NERVE FIBULAR NERVE
38. LATERAL SIDE THYROID GLAND MESENTERIC ARTERY
39. ROOT OF LATERAL VIEW SUPPLY TO
40. FACIAL NERVE DEEP FASCIA INFERIOR TO
41. KNEE JOINT ULNAR NERVE HIP JOINT
42. SUPPLY TO INFERIOR VENA SUPPLY TO
43. CRANIAL FOSSA DRAIN INTO SUPPLY TO
44. LATERAL WALL ATTACH TO CRANIAL FOSSA
45. POSTERIOR COMPARTMENT DRANGE INTO POSTERIOR COMPARTMENT
46. VAGUS NERVE INGUINAL RING HIP JOINT
47. LATERAL SURFACE INTERNAL JUGULAR HIP JOINT
48. CRANIAL CAVITY INTERNAL CAROTID CONNECTIVE TISSUE
49. HIP JOINT INFERIOR VENA SUPERIOR TO
50. CONNECTIVE TISSUE INTERNAL CAROTID IliAC SPINE
51. HYOID BONE SUPERIOR TO DIVIDES INTO
52. FEMORAL ARTERY SUPERIOR TO DIVIDES INTO
53. ILIAC ARTERY SUPERIOR TO DIVIDES INTO
54. DISTAL END ATTACHES TO DIVIDES INTO
55. ATTACHES TO FLEXOR DIGITORUM DIVIDES INTO
56. FEMORAL ARTERY SUPERIOR TO DIVIDES INTO
57. BLOOD SUPPLY LATERAL END DIVIDES INTO
Superior surface and inferior aspect are good examples of how two “general” words can combine to form a technical term; long head and short head (of a muscle) show how it is possible for two extremely common words to combine to form a highly technical term. (This demonstrates the importance of including general polysemous words and multiword units in any pedagogical list.) There are also several important three-word terms which should be included in any list, including inferior vena cava, central nervous system, and anterior cruciate ligament.

In addition to multiword terms, we find prepositional phrases, especially those including “of”, “to”, and “from”. These are found in positional phrases like posterior to, lateral to, and root of, as well as in linking expressions such as passes through, arises from, and drain into. The lists also throw up frequently used discourse structuring phrases, including due to, resulting in, such as, and in addition to, which learners need to be familiar with to be able to handle a variety of academic texts.

Is There a Common Vocabulary within the Domain of Anatomy?

As has been previously noted, anatomy is a subject placing an extremely large learning burden on the student. This is illustrated by the following definition, taken from Wikipedia:

Anatomy is the scientific study of the structure of living things including their systems, organs, and tissues. It includes the appearance and position of the various parts, the materials from which they are composed, their locations and their relationships with other parts.

It will, obviously, be impossible to provide learners with manageable lists of all the words they need to know. What is needed is a reductionist approach in order to identify a list of “core” vocabulary, the key words which are common throughout the complex domain of anatomy, weaving through and linking the different anatomical areas and concepts. However, as is apparent from the above definition, this is not going to be easy because of the numerous overlapping ways of classifying the body (e.g., structures, systems, parts, tissue types). Obvious candidates for a list of core vocabulary would be the major body parts and organs, and the terms of location and associated verbs, but it becomes more difficult when we try to determine which words to include beyond these main categories. Frequency, of course, will be a primary criterion, but we need a way of systematically selecting and organizing the words in order to ensure that those with the widest applicability are included, and that the word list we compile is user-friendly.

As a start, it may be helpful to look at the Foundational Model of Anatomy (FMA), developed at the University of Washington School of Medicine, and accessible at http://sig.biostr.washington.edu/projects/fm/AboutFM.html. The FMA is a computer-based knowledge source that claims to “represent a coherent body of explicit declarative knowledge about human anatomy”. The model broadly defines anatomy in the terms with which we are familiar: systems, structures, parts, appearance, and part-whole and spatial relationships. Clearly, words commonly used to label and specify these components will all have a place in our list of domain words.
However, in the FMA, consideration is also given to the less obvious "spaces, surfaces, lines, and points" that are associated with anatomical structures and parts.

A closer examination of the words in the frequency list reveals that it would indeed be useful to think in terms of these fundamental geometric concepts. In fact, the actual words surface (no. 28 in the list), line (no. 152), space (no. 202), and point (no. 409), all occur frequently in the anatomy corpus. Moreover, a large number of words which might be considered to belong to these categories can also be found occurring with high frequency in the corpus. This, coupled with the fact that nearly all of them are found in each chapter of the book, suggests that words such as these can certainly claim to be core anatomy words. Some examples are given below:

**Spaces:** space, cavity, tract, compartment, opening, area, segment, inlet, aperture, shaft, duct

**Surfaces:** surface, floor, plane, wall, layer, aspect, base, fascia, side, skin, membrane

**Lines:** line, midline, groove, margin, border, canal, angle, oblique, triangle, boundary, axis, length, ridge

**Points:** point, apex, end, root, terminal, node, plexus, junction, network

**IMPLICATIONS FOR MATERIALS DEVELOPMENT**

A number of implications for materials development emerge from the initial corpus analysis. As noted in Davies et al. (2014), background research in the form of interviews and conversations with medical staff and students has helped point the research in a particular direction: Medical teaching staff have stressed the importance of anatomy and also want students to be able to use English for common symptoms and diseases; medical students learn a large number of English medical terms on the basis of diagrams or with English terms tagged to texts in Japanese. For us as applied linguists, a major challenge is to develop medical English materials which expose students to key medical terms embedded in English texts.

On the basis of the ideas emerging in the background research, anatomy materials were created for an intensive medical English course that took place in September 2013, prior to the creation of the corpus described in this article. They covered the following areas: the circulatory system; the urinary system; the respiratory system; the skeletal system; the digestive system. These were based on diagrams and descriptions in Chabner (2012). As a trial, the respiratory system material was analyzed, including the following gap-filling exercise:

Air enters through the nose and travels to the ____________ (throat). From the throat, air passes through the epiglottis and ____________ (voice box) into the ____________ (windpipe). The windpipe splits into two tubes, called the ____________, that carry air into the lungs. The bronchial tubes divide into smaller tubes, called ____________, that end in the small ____________, or air sacs. The thin walls of the sacs allow oxygen to pass through them into tiny ____________ containing red blood cells. Red blood cells transport the ____________ to all parts of the body. In a similar manner, ____________ leaves the body by entering the alveoli and travelling back up to the nose.
An analysis of this text and the other respiratory system materials with the aid of the newly constructed corpus highlighted several ways in which the discourse could be improved:

1. There are no examples of the term bronchial tube in the corpus. Instead the term bronchus (plural, bronchi) is used.
2. The frequency of divides into (85 instances) is much greater than splits into (three instances), implying that where possible divide into should be used. Also, there is a further expression bifurcates into (13 instances) which can be used when something divides into two.

On the basis of this analysis the text could be re-written as follows:

Air enters through the nose and travels to the ______________ (throat). From the throat, air passes through the epiglottis and ______________ (voice box) into the ______________ (windpipe). The windpipe divides into two tubes, called the ____________, that carry air into the lungs. The bronchi further divide into smaller tubes, called ____________, that end in the small ____________, or air sacs. The thin walls of the sacs allow oxygen to pass through them into tiny ____________ containing red blood cells. Red blood cells transport the ____________ to all parts of the body. In a similar manner, ____________ leaves the body by entering the alveoli and travelling back up to the nose.

Further small issues came to light in other sections of the respiratory system materials. For example, the materials contained the term pleural space, which was included in a section on labelling a diagram. However, this does not appear in the corpus at all, while the term pleural cavity occurs 37 times. Also, the terms windpipe and voice box are used. These do not occur in the corpus, but their absence raises the important issue of the language needed in doctor-patient conversations and dialogues between medical professionals. Windpipe and voice box may be extremely useful words for doctor-patient dialogues, but not for medical papers or medical discussions.

In addition, the overall corpus analysis has highlighted the importance of such terms as anterior, posterior, superior, and inferior. These terms are much more precise than the lay terms in front of, behind, above, and below (in the same way that north, south, east, and west
are used in a geographical atlas), and the frequency with which they are used indicates that they should be introduced at an early stage in anatomy materials. In conjunction with the spaces, surfaces, lines, and points noted in the previous section, these terms would appear throughout a set of medical English teaching materials based on anatomy and common diseases/medical problems and symptoms.

It is also important to stress that corpus development is taking place in parallel with materials development. The background research for the purposes of gathering materials for a corpus of medical articles (Davies et al., 2014) led to the division of medicine into the following broad areas: Cardiovascular medicine; Digestive medicine; Respiratory medicine; Neuromusculoskeletal medicine; Infectious diseases and Immunology; Oncology; Developmental medicine; Nephrology and Endocrinology; Critical care and Anesthesiology; Sensory organology. These categories are starting to inform the planning for materials development because they also form a good way of setting up units of materials. For example, if neuromusculoskeletal medicine is divided into neurology and musculoskeletal medicine, eight categories (Cardiovascular medicine; Digestive medicine; Respiratory medicine; Neurology; Musculoskeletal medicine; Developmental medicine; Nephrology and Endocrinology; Sensory organology) can be created in which anatomical diagrams are combined with descriptions, processes and common diseases, acting as the base from which various tasks can be undertaken. The advantage of adopting this strategy is that with the development of a corpus from medical articles based on the same categories, there is a way of creating materials on reading and writing medical articles in English that fits neatly into those areas.

CONCLUSION

The investigation detailed in this paper set out to explore the merits of creating a corpus from an anatomy textbook for students, with particular reference to materials development. While there is still work to be done on tidying up parts of the corpus, it has yielded a number of interesting results and implications.

In our exploratory research we have asked a simple question: What are the best words for medical students to focus on? In seeking an answer we have not tried, as is often the case in this kind of study, to limit the words under investigation to a particular category such as “Latin words” or “semi-technical vocabulary”. Rather, we have identified the potentially most useful items as determined by their frequency in the corpus and then investigated their characteristics: the kind of words they are, how they behave in specific contexts and interrelate with other words, and how the different categories overlap. In addition, with materials development and corpus development running in parallel, the questions arising from the corpus research and the materials development have led to the construction of joint categories that will allow corpus analysis to inform materials development and vice versa: The categories that will be used in constructing an article-based corpus will be used to build units of materials; the anatomy corpus will be one of the tools for building English texts that incorporate the most frequently occurring items. In relation to this, domain anatomical terms associated with the geography and geometry
of the body will weave across and link units of material.

In terms of the future direction of the project, there is a need to use the same process to
develop a corpus relating to the diagnosis and treatment of diseases. On the basis of the
background interviews, the most suitable text for analysis is *Harrison’s Principles of Internal
Medicine*. The resulting corpus may also yield categories of regularly occurring terms that can
help to interlink different units of materials.

A final point lies with the scale of the corpus analysis. The project is going to involve three
corpora, one constructed from a textbook of anatomy, one constructed from a textbook of
internal medicine, and one from approximately 120 medical journal articles. A possible criticism
is the small size of the corpora, but we would argue that bigger is not necessarily better when
dealing with a specialized corpus; good design is the key, and our corpora and materials are
being custom-built with specific reference to a set of needs of Japanese university students. The
medical practitioners who teach these students know which medical areas, textbooks, and
articles are of high value. If those medical articles and texts can be identified, then the corpora
will be embedded within the key fields that are taught to the medical students, and given that
medicine is a profession, we believe that they will have value to other medical schools across
Japan and beyond.

**REFERENCES**

Anthony, L. (2011). AntConc (Version 3.2.4) [Computer Software]. Tokyo, Japan: Waseda
University. Available from http://www.antlab.sci.waseda.ac.jp/


Davies, W., Fraser, S., & Tatsukawa, K. (2014). A background study for the development of
medical English corpora, word lists and university course materials in Japan. *Hiroshima

Davies, W., Fraser, S., Lauer, J. & Howell, P. (2013). English for Medical Purposes: Teaching an
intensive English course to third-year medical students. *Hiroshima Studies in Language and

Livingstone.

Fraser, S. (2005). The lexical characteristics of specialized texts. In K. Bradford-Watts, C. Ikeguchi,

— 129 —


**APPENDIX 1**

The 500 Most Frequent Words in the Anatomy Corpus

1. THE
2. OF
3. AND
4. TO
5. IS
6. A
7. IN
8. FIG
9. NERVE
10. FROM
11. ARTERY
12. ANTERIOR
13. POSTERIOR
14. ARE
15. WITH
16. BY
17. MUSCLE
18. LATERAL
19. WHICH
20. ON
21. IT
22. SUPERIOR
23. MUSCLES
24. INFERIOR
25. MEDIAL
26. INTO
27. THROUGH
28. SURFACE
29. THAT
30. PART
31. AS
32. BETWEEN
33. THIS
34. BONE
35. FOR
36. AT
37. LEFT
38. WALL
39. RIGHT
40. NERVES
41. VEIN
42. BRANCHES
43. CAVITY
44. DEEP
45. BE
46. OR
47. AN
48. JOINT
49. HEAD
50. SIDE
51. LIGAMENT
52. BRANCH
53. TWO
54. NECK
55. UPPER
56. PASSES
57. LOWER
193. THYROID 238. PATIENTS 283. MESENTERIC
194. MIDLINE 239. PLANE 284. CAVA
195. SPACE 240. SINUS 285. VENA
196. SYMPATHETIC 241. TIBIAL 286. BREVIS
197. FIRST 242. ULNAR 287. RAMI
198. OBТURATOR 243. JUGULAR 288. SOFT
199. ARM 244. VISCERAL 289. VAGUS
200. THEN 245. HEART 290. BOTH
201. VERTEBRA 246. TONGUE 291. COSTAL
202. BELOW 247. BUT 292. OCCUR
203. USUALLY 248. FUNCTION 293. FIBULAR
204. CLINIC 249. PUBIC 294. IMPORTANT
205. HAND 250. SHAPED 295. LEVATOR
206. ORIGIN 251. ABDOMEN 296. AORTIC
207. PASSING 252. EAR 297. LIES
208. LUMBAR 253. MORE 298. MEDIASTINUM
209. ORBITAL 254. TUBERCLE 299. RAMUS
210. TISSUE 255. PHARYNGEAL 300. FORWARD
211. LESSER 256. LIVER 301. GLANDS
212. SURFACES 257. OPENING 302. RECTUS
213. PTERYGOID 258. PLANTAR 303. WOMEN
214. DURING 259. RING 304. DRAIN
215. FLOOR 260. ANAL 305. SUPPLIES
216. THEIR 261. FEMUR 306. TUBE
217. ORBIT 262. MEDIAN 307. MEN
218. INTERCOSTAL 263. PULMONARY 308. LARYNGEAL
219. SCIATIC 264. BONY 309. MAINLY
220. SUPERIORLY 265. SCAPULA 310. PARASYMPATHETIC
221. PROCESSES 266. SOME 311. AFTER
222. TABLE 267. CONSISTS 312. WALLS
223. ARTICULAR 268. CUTANEOUS 313. LUNG
224. DUCT 269. RELATED 314. THAN
225. AXILLARY 270. ESOPHAGUS 315. COLON
226. HAVE 271. EXTENDS 316. PENIS
227. SKULL 272. SPINE 317. SECOND
228. OBLIQUE 273. PALATE 318. ADDUCTOR
229. RIBS 274. MAXILLARY 319. BLADDER
230. PALATINE 275. THIRD 320. COLUMN
231. VENOUS 276. CONTINUES 321. FOUR
232. MANDIBLE 277. OCCIPITAL 322. TENDONS
233. MEDially 278. FORMS 323. IMMEDIATELY
234. HUMERUS 279. CREST 324. PERINEA
235. PAIN 280. DIGITORUM 325. TUBEROSITY
236. GANGLION 281. ROOT 326. DIRECTLY
237. LONG 282. MOTOR 327. INTEROSSEOUS
328. TOGETHER
329. VISCERA
330. FISSURE
331. ONTO
332. AXILLA
333. DESCENDS
334. LYMPH
335. SACRAL
336. GROOVE
337. SUBCLAVIAN
338. MANDIBULAR
339. RENAL
340. CONTAINS
341. FEMORIS
342. HIP
343. INLET
344. LACRIMAL
345. ORIGINATE
346. DRAINAGE
347. CORONARY
348. EITHER
349. LYMPHATIC
350. FIBROUS
351. TIBIA
352. APEX
353. WRIST
354. FRONTAL
355. GROUP
356. ONLY
357. CELLS
358. RADIUS
359. TEETH
360. USED
361. CAVITIES
362. CENTRAL
363. SPHENOID
364. IF
365. NEAR
366. THUMB
367. ARTERIAL
368. CONTINUOUS
369. NORMAL
370. OPTEN
371. STERNUM
372. CARDIAC
373. OUT
374. THORAX
375. ACROSS
376. CASE
377. CLINICAL
378. EXTENSION
379. INJURY
380. POPLITEAL
381. ANGLE
382. HALF
383. LARYNX
384. MARGINS
385. MINOR
386. COURSE
387. ENDS
388. STRUCTURE
389. ATTACH
390. BRAIN
391. ELBOW
392. MOVEMENT
393. NOTCH
394. PERITONEUM
395. ADDITION
396. POINT
397. SYNOVIAL
398. HYOID
399. SINUSES
400. GREAT
401. INNERVATES
402. LIKE
403. POLLICIS
404. ALVEOLAR
405. KIDNEY
406. NUMBER
407. PHARYNX
408. EYEBALL
409. FACET
410. PALMAR
411. APERTURE
412. BRACHII
413. URETHRA
414. PRESSURE
415. ZYGOMATIC
416. ARISES
417. SPINOUS
418. STOMACH
419. THEREFORE
420. TUMOR
421. CLAVICLE
422. DISEASE
423. ETHMOIDAL
424. FACE
425. ROOF
426. ASCENDING
427. ATTACHES
428. FOLD
429. OCCURS
430. TISSUES
431. ULNA
432. ANKLE
433. BOWEL
434. FRACTURE
435. MATER
436. SURGICAL
437. PARIETAL
438. SHAFT
439. ANATOMICAL
440. INNERVATE
441. PLATE
442. SHORT
443. ATRIUM
444. OPHTHALMIC
445. TERMINAL
446. DIVIDES
447. MOVEMENTS
448. TRIANGULAR
449. CROSSES
450. FINGERS
451. SKELETAL
452. TRACT
453. CARPI
454. DISC
455. PAROTID
456. SAC
457. TYMPANIC
458. BEFORE
459. CHEST
460. CONNECTIVE
461. SEGMENTS
462. UNDER
APPENDIX 2
The 300 Most Frequent Two-word Terms in the Anatomy Corpus

<table>
<thead>
<tr>
<th>Term</th>
<th>Term</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUCH</td>
<td>463.</td>
<td>476.</td>
</tr>
<tr>
<td>CLITORIS</td>
<td>464.</td>
<td>477.</td>
</tr>
<tr>
<td>JOIN</td>
<td>465.</td>
<td>478.</td>
</tr>
<tr>
<td>RECTUM</td>
<td>466.</td>
<td>479.</td>
</tr>
<tr>
<td>TRUNKS</td>
<td>467.</td>
<td>480.</td>
</tr>
<tr>
<td>ANY</td>
<td>468.</td>
<td>481.</td>
</tr>
<tr>
<td>LEAVES</td>
<td>469.</td>
<td>482.</td>
</tr>
<tr>
<td>ROOTS</td>
<td>470.</td>
<td>483.</td>
</tr>
<tr>
<td>SACRUM</td>
<td>471.</td>
<td>484.</td>
</tr>
<tr>
<td>APONEUROSIS</td>
<td>472.</td>
<td>485.</td>
</tr>
<tr>
<td>EXAMINATION</td>
<td>473.</td>
<td>486.</td>
</tr>
<tr>
<td>PRODUCE</td>
<td>474.</td>
<td>487.</td>
</tr>
<tr>
<td>SEPTUM</td>
<td>475.</td>
<td>488.</td>
</tr>
<tr>
<td>1.</td>
<td>ABDOMINAL WALL</td>
<td>28.</td>
</tr>
<tr>
<td>2.</td>
<td>CAROTID ARTERY</td>
<td>29.</td>
</tr>
<tr>
<td>3.</td>
<td>VENA CAVA</td>
<td>30.</td>
</tr>
<tr>
<td>4.</td>
<td>PELVIC CAVITY</td>
<td>31.</td>
</tr>
<tr>
<td>5.</td>
<td>POSTERIOR SURFACE</td>
<td>32.</td>
</tr>
<tr>
<td>6.</td>
<td>MEDIAL SIDE</td>
<td>33.</td>
</tr>
<tr>
<td>7.</td>
<td>ORAL CAVITY</td>
<td>34.</td>
</tr>
<tr>
<td>8.</td>
<td>Vertebral Column</td>
<td>35.</td>
</tr>
<tr>
<td>9.</td>
<td>LOWER LIMB</td>
<td>36.</td>
</tr>
<tr>
<td>10.</td>
<td>TEMPORAL BONE</td>
<td>37.</td>
</tr>
<tr>
<td>11.</td>
<td>ANTERIOR SURFACE</td>
<td>38.</td>
</tr>
<tr>
<td>12.</td>
<td>THORACIC WALL</td>
<td>39.</td>
</tr>
<tr>
<td>13.</td>
<td>GLUTEAL REGION</td>
<td>40.</td>
</tr>
<tr>
<td>14.</td>
<td>MESENTERIC ARTERY</td>
<td>41.</td>
</tr>
<tr>
<td>15.</td>
<td>UPPER LIMB</td>
<td>42.</td>
</tr>
<tr>
<td>16.</td>
<td>NASAL CAVITY</td>
<td>43.</td>
</tr>
<tr>
<td>17.</td>
<td>JUGULAR VEIN</td>
<td>44.</td>
</tr>
<tr>
<td>18.</td>
<td>SCIATIC FORAMEN</td>
<td>45.</td>
</tr>
<tr>
<td>19.</td>
<td>LATERAL SIDE</td>
<td>46.</td>
</tr>
<tr>
<td>20.</td>
<td>FACIAL NERVE</td>
<td>47.</td>
</tr>
<tr>
<td>21.</td>
<td>KNEE JOINT</td>
<td>48.</td>
</tr>
<tr>
<td>22.</td>
<td>CRANIAL FOSSA</td>
<td>49.</td>
</tr>
<tr>
<td>23.</td>
<td>LATERAL WALL</td>
<td>50.</td>
</tr>
<tr>
<td>24.</td>
<td>POSTERIOR COMPARTMENT</td>
<td>51.</td>
</tr>
<tr>
<td>25.</td>
<td>VAGUS NERVE</td>
<td>52.</td>
</tr>
<tr>
<td>26.</td>
<td>LATERAL SURFACE</td>
<td>53.</td>
</tr>
<tr>
<td>27.</td>
<td>CRANIAL CAVITY</td>
<td>54.</td>
</tr>
</tbody>
</table>
APPENDIX 3

The 100 Most Frequent Three-word Terms in the Anatomy Corpus

1. INFERIOR VENA CAVA
2. ANTERIOR ABDOMINAL WALL
3. SUPERIOR ILIAC SPINE
4. GREATER SCIATIC FORAMEN
5. SUPERIOR MESENTERIC ARTERY
6. INTERNAL JUGULAR VEIN
7. MIDDLE CRANIAL FOSSA
8. INTERNAL CAROTID ARTERY
9. POSTERIOR ABDOMINAL WALL
10. COMMON CAROTID ARTERY
11. SUPERIOR VENA CAVA
12. EXTERNAL CAROTID ARTERY
13. SPINAL CORD LEVELS
14. POSTERIOR CRANIAL FOSSA
15. INFERIOR MESENTERIC ARTERY
16. FLEXOR CARPI ULNARIS
17. INFERIOR VENA CAVA
18. SPINAL CORD LEVEL
19. MIDDLE MENINGEAL ARTERY
20. POSTERIOR SCAPULAR REGION
21. SUPERFICIAL INGUINAL RING
22. CIRCUMFLEX HUMERAL ARTERY
23. LESSER SCIATIC FORAMEN
24. BICEPS BRACHII MUSCLE
25. ANTERIOR TRIANGLE OF
26. DEEP FIBULAR NERVE
27. TENSOR VELI PALATINI
28. ANTERIOR SCALENE MUSCLE
29. DEEP PERINEAL POUCH
30. EXTERNAL ACOUSTIC MEATUS
31. FLEXOR DIGITORUM LONGUS
32. FLEXOR DIGITORUM PROFUNDUS
33. MANUBRIUM OF STERNUM
34. RAMUS OF MANDIBLE
35. TRICEPS BRACHII MUSCLE
36. COMMON FIBULAR NERVE
37. FORAMEN INFERIOR TO
38. ANTERIOR CRUCIATE LIGAMENT
39. DEEP CERVICAL NODES
40. LEVATOR VELI PALATINI
41. POSTGANGLIONIC SYMPATHETIC FIBERS
42. SUPERIOR ORBITAL FISSURE
43. DEEP PERINEAL POUCH
44. DORSALIS PEDIS ARTERY
45. HEAD OF BICEPS
46. INFERIOR ALVEOLAR NERVE
47. LEFT CORONARY ARTERY
48. SUPERIOR ALVEOLAR NERVE
49. EXTENSOR CARPI RADIALIS
50. EXTENSOR DIGITORUM LONGUS
51. GREAT APHENOUS VEIN
52. GREATER PETROSAL NERVE
53. LEVATOR PALPEBRAE SUPERIOSIS
54. OBTURATOR INTERNUS MUSCLE
55. PREGANGLIONIC PARASYMPATHETIC FIBERS
56. SCIATIC FORAMEN INFERIOR
57. INFERIOR ORBITAL FISSURE
58. SUPERIOR MESENTERIC ARTERY
59. TERES MAJOR MUSCLE
60. DEEP INGUINAL RING
61. INFERIOR THORACIC APERTURE
62. LEVATOR ANI MUSCLES
63. MAGNETIC RESONANCE IMAGE
64. ANTERIOR SUPERIOR ILIAC
65. COMMON TENDINOUS RING
66. EXTERNAL ILIAC ARTERY
67. POSTERIOR CIRCUMFLEX HUMERAL
68. POSTERIOR TIBIAL ARTERY
69. PSOAS MAJOR MUSCLE
70. WEIGHTED MAGNETIC RESONANCE
71. ANTERIOR THORACIC WALL
72. ANTERIOR TIBIAL ARTERY
73. BICEPS FEMORIS MUSCLE
74. FLEXOR CARPI RADIALIS
75. FLEXOR HALLUCIS LONGUS
76. POSTERIOR CUTANEOUS NERVE
77. SUPERIOR VENA CAVA
78. THORACIC SPINAL NERVES
79. ABDUCTOR POLLICIS LONGUS
80. FLEXOR DIGITORUM BREVIS
81. INTERNAL CAROTID ARTERY
82. LESSER PETROSAL NERVE
83. POSTERIOR ABDOMINAL REGION
84. SUPERFICIAL TEMPORAL ARTERY
85. VELI PALATINI MUSCLE
86. VISCERAL AFFERENT FIBERS
87. ABDOMINAL PREVERTEBRAL PLEXUS
88. ABDUCTOR DIGITI MINIMI
89. ASCENDING PHARYNGEAL ARTERY
90. CARPI RADIALIS BREVIS
91. CARPI RADIALIS LONGUS
92. CENTRAL NERVOUS SYSTEM
93. COMMON CAROTID ARTERY
94. COMMON FIBULAR NERVE
95. EXTENSOR HALLUCIS LONGUS
96. EXTENSOR POLLICIS LONGUS
97. INFERIOR ILIAC SPINE
98. INFRA ORBITAL NERVE
99. INTERNAL ACOUTIC MEATUS
100. ISCHIO ANAL FOSSAE
要約

解剖学教科書のコーパス分析
-基礎的研究と英語教材開発への示唆-

サイモン・フレーザー
デービス・ウォルター
達川奎三
広島大学外国語教育研究センター

本論文は、医学生を対象とした教科書Gray's Anatomy For Studentsを使い、コーパス及び英語語彙リストの作成を報告するものである。また、本研究は精選された医学研究論文のコーパスに基づき、医学生を対象とした語彙に準拠したシラバスを開発することを目指したプロジェクトの一部でもある。医学教授陣からのフィードバックでは、医学知識を下支えする専門教科、とりわけ解剖学の十分な理解をすることがまずは重要であると強調された。それ故、医学生が最も重要な専門術語を習得し、それらの使い方を学習することを支援する教材が初期段階で必要である。

解剖学コーパスの頻度に基づいて決定された、最も有益であると考えられる語彙を明らかにし、これらの語彙の特徴を探求した。当然のことながら、名詞の頻度が非常に高いことが分ったが、加えて形容詞、動詞、そして空間的（位置）関係を描写する前置詞も多かった。語彙がどのように組み合わされるかは、専門性の見地から特に重要であることが示され、単一の語彙単位を越える繋がりを無視できないことが判明した。加えて、頻度リストでの語彙をよく吟味してみると、解剖学における共通の、または核となる語彙を特定することが可能であることが示唆された。それらは、解剖学のすべての範疇で使われる、身体の位置や形状に関わる専門術語である。解剖学コーパスの分析はまだ十分なものではないが、本基礎的研究で判明したことは、医学英語教材を開発するための多くの示唆を含み、これらの示唆について詳細に議論する。