# 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 1. Top 100 Words in the Anatomy Corpus (Frequency)**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1. | THE | (46419) | 33. | THIS | (1097) | 65. | FOSSA | (611) |
| 2. | OF | (18400) | 34. | BONE | (1037) | 66. | ABDOMINAL | (604) |
| 3. | AND | (13328) | 35. | FOR | (1014) | 67. | BODY | (595) |
| 4. | TO | (6964) | 36. | AT | (1002) | 68. | CAN | (594) |
| 5. | IS | (5383) | 37. | LEFT | (967) | 69. | FIBERS | (591) |
| 6. | A | (5112) | 38. | WALL | (965) | 70. | REGION | (590) |
| 7. | IN | (5030) | 39. | RIGHT | (961) | 71. | VESSELS | (566) |
| 8. | FIG | (3132) | 40. | NERVES | (943) | 72. | THORACIC | (560) |
| 9. | NERVE | (2974) | 41. | VEIN | (921) | 73. | ALSO | (548) |
| 10. | FROM | (2792) | 42. | BRANCHES | (920) | 74. | SPINAL | (548) |
| 11. | ARTERY | (2379) | 43. | CAVITY | (881) | 75. | SMALL | (540) |
| 12. | ANTERIOR | (2325) | 44. | DEEP | (862) | 76. | INTERNAL | (539) |
| 13. | POSTERIOR | (2280) | 45. | BE | (850) | 77. | MAJOR | (532) |
| 14. | ARE | (2195) | 46. | OR | (850) | 78. | MIDDLE | (531) |
| 15. | WITH | (2024) | 47. | AN | (841) | 79. | FORAMEN | (527) |
| 16. | BY | (1995) | 48. | JOINT | (829) | 80. | EXTERNAL | (507) |
| 17. | MUSCLE | (1922) | 49. | HEAD | (817) | 81. | WAS | (505) |
| 18. | LATERAL | (1862) | 50. | SIDE | (808) | 82. | PROCESS | (499) |
| 19. | WHICH | (1753) | 51. | LIGAMENT | (806) | 83. | MARGIN | (485) |
| 20. | ON | (1722) | 52. | BRANCH | (792) | 84. | VERTEBRAL | (482) |
| 21. | IT | (1720) | 53. | TWO | (762) | 85. | PASS | (478) |
| 22. | SUPERIOR | (1604) | 54. | NECK | (745) | 86. | SUPPLY | (478) |
| 23. | MUSCLES | (1594) | 55. | UPPER | (729) | 87. | CANAL | (476) |
| 24. | INFERIOR | (1512) | 56. | PASSES | (727) | 88. | FASCIA | (473) |
| 25. | MEDIAL | (1372) | 57. | LOWER | (695) | 89. | PLEXUS | (469) |
| 26. | INTO | (1360) | 58. | THESE | (691) | 90. | SUPERFICIAL | (464) |
| 27. | THROUGH | (1278) | 59. | VEINS | (673) | 91. | BLOOD | (463) |
| 28. | SURFACE | (1231) | 60. | PELVIC | (663) | 92. | VIEW | (455) |
| 29. | THAT | (1215) | 61. | MAY | (657) | 93. | TENDON | (453) |
| 30. | PART | (1196) | 62. | ITS | (651) | 94. | STRUCTURES | (436) |
| 31. | AS | (1163) | 63. | EACH | (644) | 95. | OTHER | (433) |
| 32. | BETWEEN | (1110) | 64. | ARTERIES | (621) | 96. | THEY | (431) |

| | | | | | |
|---|---|---|---|---|---|
| 97. | ASSOCIATED | (426) | 99. | LIMB | (420) |
| 98. | PATIENT | (423) | 100. | ANATOMY | (419) |

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.)

**TABLE 2. Top 100 Two-word Combinations in the Anatomy Corpus**

| | | | |
|---|---|---|---|
| 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
16. VENA CAVA
17. PASSES THROUGH
18. ATTACHED TO
19. PELVIC CAVITY
20. POSTERIOR SURFACE
21. PASS THROUGH
22. LATERAL TO
23. MEDIAL SIDE
24. ORAL CAVITY
25. VERTEBRAL COLUMN
26. ATTACHMENT OF
27. LOWER LIMB
28. TEMPORAL BONE
29. ANTERIOR SURFACE
30. THORACIC WALL
31. GLUTEAL REGION
32. MESENTERIC ARTERY
33. UPPER LIMB
34. NASAL CAVITY
35. BRANCHES FROM
36. JUGULAR VEIN
37. SCIATIC FORAMEN
38. LATERAL SIDE
39. ROOT OF
40. FACIAL NERVE
41. KNEE JOINT
42. SUPPLY TO
43. CRANIAL FOSSA
44. LATERAL WALL
45. POSTERIOR COMPARTMENT
46. VAGUS NERVE
47. LATERAL SURFACE
48. CRANIAL CAVITY
49. HIP JOINT
50. CONNECTIVE TISSUE
51. HYOID BONE
52. ANTERIOR RAMI
53. SUPERIOR TO
54. ILIAC SPINE
55. DIVIDES INTO
56. ARISES FROM
57. BLOOD SUPPLY
58. DISTAL END
59. ATTACHES TO
60. FLEXOR DIGITORUM
61. ILIAC ARTERY
62. FEMORAL ARTERY
63. PASSING THROUGH
64. INGUINAL LIGAMENT
65. MEDIAN NERVE
66. THYROID GLAND
67. ANTERIOR VIEW
68. EDGE OF
69. INFERIOR VENA
70. FIBULAR NERVE
71. LIGAMENT OF
72. PORTION OF
73. SUPERIOR MESENTERIC
74. ANTERIOR ABDOMINAL
75. SOFT PALATE
76. MIDDLE EAR
77. RADIAL NERVE
78. SUBCLAVIAN ARTERY
79. DEEP FASCIA
80. ULNAR NERVE
81. LATERAL VIEW
82. MAJOR MUSCLE
83. ABDOMINAL AORTA
84. INFRATEMPORAL AORTA
85. INFRATEMPORAL FOSSA
86. INTERNAL CAROTID
87. ATTACH TO
88. DRAIN INTO
89. INGUINAL RING
90. INTERNAL JUGULAR
91. MOST OF
92. SUCH AS
93. ANTERIOR COMPARTMENT
94. CUTANEOUS NERVE
95. ELBOW JOINT
96. LOWER LIMB
97. SPLANCHNIC NERVES
98. SUPERIOR ILIAC
99. TRIANGLE OF
100. GREATER SCIATIC

*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.

| | | | | |
|---|---|---|---|---|
| capillaries | bronchioles | carbon dioxide | alveoli | |
| bronchial tubes | pharynx | oxygen | trachea | larynx |

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.

| | | | | |
|---|---|---|---|---|
| capillaries | bronchioles | carbon dioxide | alveoli | |
| **bronchi** | pharynx | oxygen | trachea | larynx |

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/

Baker, M. (1988). Sub-technical vocabulary and the ESP teacher: An analysis of some rhetorical items in medical journal articles. *Reading in a Foreign Language, 4* (2), 91–105.

Chabner, D-E. (2012). *Medical Terminology* (*6th Edition*). St Louis, Missouri: Elsevier.

Chung, T. M. & Nation, I. S. P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language, 15* (2), 102–116.

Chung, T.M. & Nation, I.S.P. (2004). Identifying technical vocabulary. *System,* 32, 251–263.

Davies, W. (2013). Linking a medical practitioner's use of English for Medical Purposes in Japan to skills development at university. *Hiroshima Studies in Language and Language Education, 16,* 89–101.

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 Studies in Language and Language Education, 17,* 105–117.

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 Language Education, 16,* 49–59.

Drake, R., Vogl A.W., & Mitchell, W.M. (2004). *Gray's Anatomy for Students*. London: Churchill Livingstone.

Fraser, S. (2005). The lexical characteristics of specialized texts. In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.) *JALT 2004 Conference Proceedings*. Tokyo: JALT.

Fraser, S. (2007). Providing ESP learners with the vocabulary they need: Corpora and the creation of specialized word lists. *Hiroshima Studies in Language and Language Education, 10,* 127–143.

Fraser, S. (2009). Breaking down the divisions between general, academic, and technical vocabulary: The establishment of a single, discipline-based word list for ESP learners. *Hiroshima Studies in Language and Language Education, 12,* 151–167.

Fraser, S. (2012). Factors affecting the learning of technical vocabulary. *Hiroshima Studies in Language and Language Education, 15,* 123–142.

Fraser, S. (2013). Building corpora and compiling pedagogical lists for university medical students. *Hiroshima Studies in Language and Language Education, 16,* 65–88.

Hsu, W. (2013). Bridging the vocabulary gap for EFL medical undergraduates: The establishment of a medical word list. *Language Teaching Research, 17*(4), 454–484.

Longo, D. et al. (2012). *Harrison's Principles of Internal Medicine* (18th Edition). McGraw Hill.

Lowe, I. (2010). www.scientificlanguage.com/reviews/chung-nation.pdf (accessed on 30 October, 2013)

Moore, K.L. & Dalley, A.F. (1999). *Clinically Oriented Anatomy* (*4th Edition*). Philadelphia: Lippincott, Williams and Wilkins.

Wang, J., Liang, S-l, & Ge, G-c. (2008). Establishment of a medical academic word list. *English for Specific Purposes 27* (4), 442–458.

## APPENDIX 1

### The 500 Most Frequent Words in the Anatomy Corpus

| | | | | | |
|---|---|---|---|---|---|
| 1. | THE | 20. | ON | 39. | RIGHT |
| 2. | OF | 21. | IT | 40. | NERVES |
| 3. | AND | 22. | SUPERIOR | 41. | VEIN |
| 4. | TO | 23. | MUSCLES | 42. | BRANCHES |
| 5. | IS | 24. | INFERIOR | 43. | CAVITY |
| 6. | A | 25. | MEDIAL | 44. | DEEP |
| 7. | IN | 26. | INTO | 45. | BE |
| 8. | FIG | 27. | THROUGH | 46. | OR |
| 9. | NERVE | 28. | SURFACE | 47. | AN |
| 10. | FROM | 29. | THAT | 48. | JOINT |
| 11. | ARTERY | 30. | PART | 49. | HEAD |
| 12. | ANTERIOR | 31. | AS | 50. | SIDE |
| 13. | POSTERIOR | 32. | BETWEEN | 51. | LIGAMENT |
| 14. | ARE | 33. | THIS | 52. | BRANCH |
| 15. | WITH | 34. | BONE | 53. | TWO |
| 16. | BY | 35. | FOR | 54. | NECK |
| 17. | MUSCLE | 36. | AT | 55. | UPPER |
| 18. | LATERAL | 37. | LEFT | 56. | PASSES |
| 19. | WHICH | 38. | WALL | 57. | LOWER |

58. THESE
59. VEINS
60. PELVIC
61. MAY
62. ITS
63. EACH
64. ARTERIES
65. FOSSA
66. ABDOMINAL
67. BODY
68. CAN
69. FIBERS
70. REGION
71. VESSELS
72. THORACIC
73. ALSO
74. SPINAL
75. SMALL
76. INTERNAL
77. MAJOR
78. MIDDLE
79. FORAMEN
80. EXTERNAL
81. WAS
82. PROCESS
83. MARGIN
84. VERTEBRAL
85. PASS
86. SUPPLY
87. CANAL
88. FASCIA
89. PLEXUS
90. SUPERFICIAL
91. BLOOD
92. VIEW
93. TENDON
94. STRUCTURES
95. OTHER
96. THEY
97. ASSOCIATED
98. PATIENT
99. LIMB
100. ANATOMY
101. SKIN
102. ONE
103. MEMBRANE
104. GREATER
105. ILIAC
106. CORD
107. FEMORAL
108. MOST
109. NASAL
110. BONES
111. WHEN
112. HAS
113. THIGH
114. FORM
115. TRANSVERSE
116. TRUNK
117. END
118. FOOT
119. FLEXOR
120. AROUND
121. CERVICAL
122. ASPECT
123. DISTAL
124. BASE
125. INGUINAL
126. POSITION
127. COMMON
128. GLAND
129. THREE
130. ALONG
131. BORDER
132. NODES
133. NOT
134. CRANIAL
135. REGIONAL
136. CAROTID
137. ARCH
138. WHERE
139. ANTERIORLY
140. OVER
141. WITHIN
142. SYSTEM
143. AORTA
144. LARGE
145. LATERALLY
146. LINE
147. TEMPORAL
148. ATTACHMENT
149. PARTS
150. ADJACENT
151. CARTILAGE
152. FACIAL
153. JOINTS
154. COMPARTMENT
155. FORMED
156. GLUTEAL
157. EXTENSOR
158. DIAPHRAGM
159. DORSAL
160. PELVIS
161. TRIANGLE
162. FOREARM
163. INNERVATED
164. THERE
165. ALL
166. LAYER
167. LEVEL
168. RIB
169. INNERVATION
170. ENTER
171. ATTACHED
172. LEG
173. BECAUSE
174. JUST
175. PROXIMAL
176. LIGAMENTS
177. AREA
178. VERTEBRAE
179. ENTERS
180. INCLUDE
181. REGIONS
182. POSTERIORLY
183. INFERIORLY
184. LONGUS
185. ORAL
186. KNEE
187. ABOVE
188. RADIAL
189. SENSORY
190. BACK
191. PERINEUM
192. ORIGINATES

193. THYROID
194. MIDLINE
195. SPACE
196. SYMPATHETIC
197. FIRST
198. OBTURATOR
199. ARM
200. THEN
201. VERTEBRA
202. BELOW
203. USUALLY
204. CLINIC
205. HAND
206. ORIGIN
207. PASSING
208. LUMBAR
209. ORBITAL
210. TISSUE
211. LESSER
212. SURFACES
213. PTERYGOID
214. DURING
215. FLOOR
216. THEIR
217. ORBIT
218. INTERCOSTAL
219. SCIATIC
220. SUPERIORLY
221. PROCESSES
222. TABLE
223. ARTICULAR
224. DUCT
225. AXILLARY
226. HAVE
227. SKULL
228. OBLIQUE
229. RIBS
230. PALATINE
231. VENOUS
232. MANDIBLE
233. MEDIALLY
234. HUMERUS
235. PAIN
236. GANGLION
237. LONG
238. PATIENTS
239. PLANE
240. SINUS
241. TIBIAL
242. ULNAR
243. JUGULAR
244. VISCERAL
245. HEART
246. TONGUE
247. BUT
248. FUNCTION
249. PUBIC
250. SHAPED
251. ABDOMEN
252. EAR
253. MORE
254. TUBERCLE
255. PHARYNGEAL
256. LIVER
257. OPENING
258. PLANTAR
259. RING
260. ANAL
261. FEMUR
262. MEDIAN
263. PULMONARY
264. BONY
265. SCAPULA
266. SOME
267. CONSISTS
268. CUTANEOUS
269. RELATED
270. ESOPHAGUS
271. EXTENDS
272. SPINE
273. PALATE
274. MAXILLARY
275. THIRD
276. CONTINUES
277. OCCIPITAL
278. FORMS
279. CREST
280. DIGITORUM
281. ROOT
282. MOTOR
283. MESENTERIC
284. CAVA
285. VENA
286. BREVIS
287. RAMI
288. SOFT
289. VAGUS
290. BOTH
291. COSTAL
292. OCCUR
293. FIBULAR
294. IMPORTANT
295. LEVATOR
296. AORTIC
297. LIES
298. MEDIASTINUM
299. RAMUS
300. FORWARD
301. GLANDS
302. RECTUS
303. WOMEN
304. DRAIN
305. SUPPLIES
306. TUBE
307. MEN
308. LARYNGEAL
309. MAINLY
310. PARASYMPATHETIC
311. AFTER
312. WALLS
313. LUNG
314. THAN
315. COLON
316. PENIS
317. SECOND
318. ADDUCTOR
319. BLADDER
320. COLUMN
321. FOUR
322. TENDONS
323. IMMEDIATELY
324. PERINEA
325. TUBEROSITY
326. DIRECTLY
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. OFTEN
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

463. SUCH
464. CLITORIS
465. JOIN
466. RECTUM
467. TRUNKS
468. ANY
469. LEAVES
470. ROOTS
471. SACRUM
472. APONEUROSIS
473. EXAMINATION
474. PRODUCE
475. SEPTUM
476. TERES
477. ARCHES
478. FURTHER
479. LITTLE
480. SYMPHYSIS
481. BICEPS
482. BRACHIA
483. FLUID
484. PTERYGOPALATINE
485. SHOULDER
486. UP
487. FRACTURES
488. PLACED
489. SEE
490. VESSEL
491. FIBULA
492. GANGLIA
493. HEPATIC
494. INFRATEMPORAL
495. MENINGEAL
496. OPTIC
497. PREGANGLIONIC
498. BODIES
499. CARPAL
500. DOWN

## APPENDIX 2

### The 300 Most Frequent Two-word Terms in the Anatomy Corpus

1. ABDOMINAL WALL
2. CAROTID ARTERY
3. VENA CAVA
4. PELVIC CAVITY
5. POSTERIOR SURFACE
6. MEDIAL SIDE
7. ORAL CAVITY
8. VERTEBRAL COLUMN
9. LOWER LIMB
10. TEMPORAL BONE
11. ANTERIOR SURFACE
12. THORACIC WALL
13. GLUTEAL REGION
14. MESENTERIC ARTERY
15. UPPER LIMB
16. NASAL CAVITY
17. JUGULAR VEIN
18. SCIATIC FORAMEN
19. LATERAL SIDE
20. FACIAL NERVE
21. KNEE JOINT
22. CRANIAL FOSSA
23. LATERAL WALL
24. POSTERIOR COMPARTMENT
25. VAGUS NERVE
26. LATERAL SURFACE
27. CRANIAL CAVITY
28. HIP JOINT
29. CONNECTIVE TISSUE
30. HYOID BONE
31. ANTERIOR RAMI
32. ILIAC SPINE
33. BLOOD SUPPLY
34. DISTAL END
35. FLEXOR DIGITORUM
36. ILIAC ARTERY
37. FEMORAL ARTERY
38. INGUINAL LIGAMENT
39. MEDIAN NERVE
40. THYROID GLAND
41. ANTERIOR VIEW
42. INFERIOR VENA
43. FIBULAR NERVE
44. SUPERIOR MESENTERIC
45. ANTERIOR ABDOMINAL
46. SOFT PALATE
47. MIDDLE EAR
48. RADIAL NERVE
49. SUBCLAVIAN ARTERY
50. DEEP FASCIA
51. ULNAR NERVE
52. LATERAL VIEW
53. MAJOR MUSCLE
54. ABDOMINAL AORTA

55. INFRATEMPORAL AORTA
56. INFRATEMPORAL FOSSA
57. INTERNAL CAROTID
58. INGUINAL RING
59. INTERNAL JUGULAR
60. ANTERIOR COMPARTMENT
61. CUTANEOUS NERVE
62. ELBOW JOINT
63. LOWER LIMB
64. SPLANCHNIC NERVES
65. SUPERIOR ILIAC
66. GREATER SCIATIC
67. PARASYMPATHETIC FIBERS
68. LYMPH NODES
69. OBTURATOR INTERNUS
70. INFERIOR MARGIN
71. CORONARY ARTERY
72. MANDIBULAR NERVE
73. NASAL CAVITIES
74. POSTERIOR ABDOMINAL
75. SCIATIC NERVE
76. BLOOD VESSELS
77. SYMPATHETIC FIBERS
78. PELVIC BONE
79. VELI PALATINI
80. PAROTID GLAND
81. POSTERIOR ASPECT
82. PTERYGOPALATINE FOSSA
83. COMMON CAROTID
84. INGUINAL CANAL
85. UPPER LIMB
86. PELVIC FLOOR
87. PETROSAL NERVE
88. RIGHT ATRIUM
89. NERVOUS SYSTEM
90. INFERIOR SURFACE
91. POSTERIOR VIEW
92. THORACIC APERTURE
93. DURA MATER
94. ILIAC CREST
95. ORBITAL FISSURE
96. PUBIC SYMPHYSIS
97. SUPERFICIAL FASCIA
98. ABDOMINAL CAVITY
99. ANTERIOR SUPERIOR
100. BRANCHII MUSCLE
101. EXTERNAL CAROTID
102. CARPI ULNARIS
103. LATERAL MARGIN
104. MAXILLARY NERVE
105. ACOUSTIC MEATUS
106. DIGITORUM LONGUS
107. FACIAL NERVE
108. INTRINSIC MUSCLES
109. MIDDLE CRANIAL
110. AXILLARY ARTERY
111. INFRA ORBITAL
112. POLLICIS LONGUS
113. GLENOHUMERAL JOINT
114. PELVIC INLET
115. PREVERTEBRAL PLEXUS
116. TIBIAL ARTERY
117. INFERIOR MESENTERIC
118. INNERVATION FUNCTION
119. MEDIAL WALL
120. SUPERIOR MEDIASTINUM
121. ANAL CANAL
122. ANKLE JOINT
123. COSTAL CARTILAGE
124. FLEXOR CARPI
125. FRONTAL BONE
126. INTEROSSEOUS MEMBRANE
127. MUSCLE FIBERS
128. OCCIPITAL BONE
129. POSTERIOR SUPERIOR
130. TIBIAL NERVE
131. OCULOMOTOR NERVE
132. PHARYNGOTYMPANIC TUBE
133. POSTERIOR WALL
134. SUPERIOR ALVEOLAR
135. SUPERIOR SURFACE
136. VERTEBRAL CANAL
137. CARPI RADIALIS
138. EXTENSOR DIGITORUM
139. LONG HEAD
140. MAXILLARY ARTERY
141. SPHENOID BONE
142. SPINOUS PROCESSES
143. ARTERIAL SUPPLY
144. BICEPS BRACHII

145. PELVIC WALL
146. TRICEPS BRACHII
147. ALVEOLAR NERVE
148. ANTERIOR ASPECT
149. DEEP SURFACE
150. FIBROUS MEMBRANE
151. LARYNGEAL NERVE
152. POPLITEAL FOSSA
153. SAPHENOUS VEIN
154. SUPERIOR VENA
155. THROID CARTILAGE
156. VERTEBRAL LEVEL
157. DIGITI MINIMI
158. GREAT TOE
159. INTERNAL ILEAC
160. PERITONEAL CAVITY
161. PETROUS PART
162. TRANSVERSE PROCESSES
163. BRACHIA PLEXUS
164. FEMORIS MUSCLE
165. TRIGEMINAL NERVE
166. FLEXOR RETINACULUM
167. LEVATOR ANI
168. MENINGEAL ARTERY
169. POLLICIS BREVIS
170. POSTERIOR TRIANGLE
171. ANTERIOR PART
172. CORD LEVELS
173. GLOSSOPHARYNGEAL NERVE
174. PHARYNGEAL WALL
175. HARD PALATE
176. LONGUS MUSCLE
177. PERINEA MEMBRANE
178. PIRIFORMIS MUSCLE
179. PLEURAL CAVITY
180. SURFACE ANATOMY
181. ZYGOMATIC BONE
182. ANTERIOR RAMUS
183. ANTERIOR SUPERIOR
184. ANTERIOR WALL
185. CELIAC TRUNK
186. FEMORAL TRIANGLE
187. FORAMEN MAGNUM
188. INFERIOR GLUTEAL
189. INSERT INTO
190. JOINT CAPSULE
191. LESSER SCIATIC
192. MEDIAL BORDER
193. PTERYGOID PROCESS
194. SPINAL NERVE
195. SUBARACHNOID SPACE
196. COMMON FIBULAR
197. CRICOID CARTILAGE
198. CRUCIATE LIGAMENT
199. CUBITAL FOSSA
200. FACIAL ARTERY
201. FEMORAL NERVE
202. INTERCOSTAL SPACE
203. LATERAL PTERYGOID
204. MEDIAL MARGIN
205. POSTERIOR CRANIAL
206. PULMONARY ARTERY
207. SCAPULAR REGION
208. TYMPANIC MEMBRANE
209. UROGENITAL TRIANGLE
210. VERTEBRAL BODIES
211. BLOOD FLOW
212. BRACHIOCEPHALIC VEIN
213. HALLUCIS LONGUS
214. PECTORALIS MAJOR
215. PLANTAR NERVE
216. POSTERIOR MARGIN
217. SIGMOID COLON
218. SUPERIOR MARGIN
219. SYMPATHETIC TRUNK
220. ANTERIOR TRIANGLE
221. FEMORAL VEIN
222. HORIZONTAL PLANE
223. OBLIQUE MUSCLE
224. OPTIC NERVE
225. QUADRATUS FEMORIS
226. TEMPORAL FOSSA
227. CAROTID ARTERIES
228. GLUTEAL NERVE
229. LEFT VENTRICLE
230. LINGUAL NERVE
231. MEDIAL MALLEOLUS
232. PERINEA POUCH
233. RADIAL ARTERY
234. SCALENE MUSCLE

235. SUPERFICIAL INGUINAL
236. VAGUS NERVES
237. CT SCAN
238. DEEP CERVICAL
239. LATERAL ASPECT
240. LUMBAR VERTEBRAE
241. OPHTHALMIC NERVE
242. PHRENIC NERVE
243. RIGHT VENTRICLE
244. SOFT TISSUE
245. SOFT TISSUES
246. SYNOVIAL MEMBRANE
247. BACK MUSCLES
248. BICEPS FEMORIS
249. BLOOD PRESSURE
250. BRACHIA ARTERY
251. DISTAL FOREARM
252. HEPATIC ARTERY
253. INFERIOR VENA
254. INTERNAL THORACIC
255. ISCHIAL TUBEROSITY
256. LEFT ATRIUM
257. MEDIAL PTERYGOID
258. POSTERIOR BORDER
259. SENSORY FIBERS
260. SUPERIOR PART
261. SUPERIOR VIEW
262. AMOUNT OF
263. ANAL TRIANGLE
264. ANATOMICAL POSITION
265. COLIC ARTERY
266. CORD LEVEL
267. EXTERNAL ILIAC
268. GREATER PALATINE
269. INFERIOR ASPECT
270. LATERAL BORDER
271. LYMPHATIC DRAINAGE
272. MIDDLE MENINGEAL
273. PANCREATICODUODENAL ARTERY
274. POSTERIOR SCAPULAR
275. PTERYGOID CANAL
276. SMOOTH MUSCLE
277. WRIST JOINT
278. ARTERIAL SUPPLY
279. CIRCUMFLEX HUMERAL
280. CLINICAL CASES
281. DEEP BRANCH
282. GREATER WING
283. HUMERAL ARTERY
284. INTERVERTEBRAL DISC
285. JUGULAR FORAMEN
286. LOWER BORDER
287. LYMPHATIC VESSELS
288. MEDIAL ASPECT
289. PORTAL VEIN
290. POSTERIOR PART
291. RADIAL NERVE
292. SPINOUS PROCESS
293. TEMPORAL ARTERY
294. THORACIC ARTERY
295. TRANSVERSE PROCESS
296. ULNAR ARTERY
297. VERTEBRAL BODY
298. ARTICULAR SURFACES
299. BRACHIAL PLEXUS
300. CARDIAC VEIN
301. CARPAL BONES
302. CAVERNOUS SINOUS
303. COLLATERAL LIGAMENT
304. EXTERNAL OBLIQUE
305. GASTROINTESTINAL TRACT
306. INFERIOR ALVEOLAR
307. ISCHIAL SPINE
308. LACRIMAL GLAND
309. LONGUS TENDON
310. MEDIAN NERVE
311. NERVE ENTERS
312. PELVIC VISCERA
313. SUPERIOR BORDER
314. THORACIC AORTA
315. ANTERIOR SCALENE
316. CARPAL TUNNEL
317. CERVICAL NODES
318. COSTAL CARTILAGES
319. DIGITORUM PROFUNDUS
320. DORSAL ASPECT
321. FORAMEN OVALE
322. GREATER TROCHANTER
323. MAGNETIC RESONANCE
324. OBTURATOR NERVE

325. PALPEBRAE SUPERIOSIS
326. POSTERIOR INTERCOSTAL
327. PROXIMAL PHALANX
328. SPINAL SEGMENTS
329. SUPERIOR MESENTERIC
330. TERMINAL BRANCHES
331. THORACIC DUCT
332. THORACIC VERTEBRAE
333. VENOUS DRAINAGE
334. VOCAL FOLDS
335. ACCESSORY NERVE
336. ANTERIOR BORDER
337. ANTERIOR VIEW
338. ARTICULAR SURFACE
339. CRANIAL NERVES
340. DEEP FIBULAR
341. GLUTEAL ARTERY
342. PREGANGLIONIC PARASYMPATHETIC
343. PULMONARY TRUNK
344. RIGHT LUNG
345. SPERMATIC CORD
346. STERNOCLEIDOMASTOID MUSCLE
347. ADDUCTOR MAGNUS
348. CERVICAL NERVES
349. DIGASTRIC MUSCLE
350. EXTENSOR CARPI
351. EXTERNAL ACOUSTIC
352. FIBULARIS LONGUS
353. GLUTEUS MAXIMUS

## 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 ACOUSTIC 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。