Identifying technical vocabulary. An assessment of the views of TM Chung and P Nation (2003, 2004)

First published 5 January 2010

A. Introduction

In two largely similar articles Chung and Nation set out a four point rating scale for distinguishing technical from non-technical language, and compare it with three other methods: a dictionary, textual clues, and computer programs. I have summarised many other viewpoints at length in the article on terminology, www.scientificlanguage.com/esp/terminology.pdf. Here I wish to interact with these authors in detail.

B. Discussion

1. Factual error: Bone marrow is not a synonym for hematopoietic tissue

Although one of the authors is a nurse and teaches public health, there is a major factual error. The following example is given of defining through the use of synonyms in brackets.

> "bone marrow (hematopoietic tissue) that forms blood cells" (2004:256).

These terms are NOT synonymous. Bone marrow is a type of hematopoietic tissue as this definition explains:

> " pertaining to the formation of blood or blood cells; hemopoietic stem cells in bone marrow" wordnetweb.princeton.edu/perl/webwn (accessed 23 December 2009).

Typing 'bone marrow' into http://www.medilexicon.com/medicaldictionary.php gives the synonym *medulla ossium*.

In fact, even without recognising the word 'hematopoietic' my suspicions were raised, simply because the linguistic clues suggest that class and sub-class is involved. In addition, I cannot recall 'bone marrow' being classified as a tissue. Chung & Nation and others when using bracketed terms as an indication that a definition or a synonym is involved need to take care: brackets can also indicate explanation or commentary.

The online version of the huge and well known textbook, Harrison's Internal medicine has:

Section 2. Hematopoietic Disorders

Chapter 98 Iron Deficiency and Other Hypoproliferative Anemias
Chapter 99 Disorders of Hemoglobin
Latest Update: 06/30/08: Hydroxyurea Improves the Clinical Course in Adults with Sickle Cell Anemia
Chapter 100 Megaloblastic Anemias
Chapter 101 Hemolytic Anemias and Anemia Due to Acute Blood Loss
Chapter 102 Aplastic Anemia, Myelodysplasia, and Related Bone Marrow Failure Syndromes
Chapter 103 Polycythemia Vera and Other Myeloproliferative Diseases
Chapter 104 Acute and Chronic Myeloid Leukemia
Chapter 105 Malignancies of Lymphoid Cells
Chapter 106 Plasma Cell Disorders
Chapter 107 Transfusion Biology and Therapy
Latest Update: 07/03/08: Young Blood: Does It Matter?
Chapter 108 Hematopoietic Cell Transplantation

2. Technical words are those words which are rare outside the speciality

This is a good idea. The authors try to identify a technical word as a word that does not exist, or exists rarely in a representative corpus excluding the medical section (2004:259). The cutoff point was set at 50. Thus, the technical words in medicine were either non-existent in a general corpus, or occurred 50 times less frequently.

This assumes that all other specialities are fully represented, which would mean a corpus much bigger than that which was used. The authors used the LOB corpus (1978) combined with the Wellington corpus of Written English (1993) both of which used 500 texts of 2000 words each, ie 1 million words each.

At first sight this appears to be incredibly small. Just considering medicine, the spelling checkers and dictionaries dwarf the corpus used for comparison.

Commercial specialised Spelling checkers and other huge sources
http://www.spellex.com/products/med.htm sells spelling checkers for medical specialities. A quick look showed them offering the biotechnology dictionary of 130,000 terms (presumably tokens) and a combined medical and pharmaceutical package of 450,000 terms.

http://www.interfold.com/translator/medlinks.htm lists hundreds of links to medical glossaries. http://www.medilexicon.com/ links to a dictionary of over 200,000 medical abbreviations, Stedman's Electronic Medical Dictionary has more than 107,000 terms. This spelling checker provides nearly half a million medical, pharmaceutical and bioscience terms from over 60 major medical specialities http://www.stedmans.com/product.cfm/611/228 with the option of paying for quarterly updates.

And that is just one subject: medicine! I have not covered physics, chemistry, botany, the classification of species and so on. Perhaps a useful comparison would be to ask if the word is found in the specialist spelling checkers or the general spelling checkers. The problem with this is that the general spelling checkers will include the more well known, the more frequent technical words.

Every speciality - Sciences and Humanities, has its own technical words. I do not know how large the list of these words is. What can be said with confidence, is that in a balanced corpus of one million words, very few will be present. That is I think what Chung and Nation are assuming.

Given the number of technical words, it can be safely assumed that some of them will be used frequently, and some of them rarely, with all possibilities in between. Therefore I conclude that the idea is a good one. The smallness of the corpus is perhaps its strength, in that it prioritises the common words.

3. **"A considerable number of technical words were from the first 2000 words of English and the Academic Word list"** (2003: Abstract)
This conclusion does not surprise me. It would be interesting to find out how many words from these lists actually have a specific technical sense. In my definition, such words would be called 'semi-technical'. Unfortunately, the authors never use this term or something like it. The problem repeats itself with the rating scales.

4. **Commentary on the rating scale**
Adapted from Table 1: A rating scale for finding technical words (as applied to the anatomy text). The original text is here, with my interpretation to the right. The italics are in the original, but the bold is mine.

Step (adapted from) http://nflrc.hawaii.edu/rfl/October2003/chung/chung.html	**Chung & Nation**	**Lowe**
1. Words such as function words that have a meaning that has no particular relationship with the field of anatomy, that is, words independent of the subject matter. Examples are: *the, is, between, it, by, 12, adjacent, amounts, common, commonly, directly, constantly, early, and especially.*	Common	Common
2. Words that have a meaning that is minimally related to the field of anatomy in that they describe the positions, movements, or features of the body. Examples are: *superior, part, forms, pairs, structures, surrounds, supports, associated, lodges, protects*.	Common	Common or semi-technical
3. Words that have a meaning that is closely related to the field of anatomy. They refer to parts, structures or functions of the body, such as the regions of the body and systems of the body. **Such words are also used in general language**. The words may have some restrictions of usage depending on the subject field. Examples are: *chest, trunk, neck, abdomen, ribs, breast, cage, cavity, shoulder, girdle, skin, muscles, wall, heart, lungs, organs, liver, bony, abdominal, breathing.* **Words in this category may be technical terms in a specific field like anatomy and yet may occur with the same meaning in other fields and not be technical terms in those fields.** "Words at Step 3 may have polysemes that occur in general use, and in some cases occur in general use with little change in meaning, for example breathe and bony. Step 4 includes words like thorax and mammary which may be known in other fields but which have a technical flavour. Even though they are used outside anatomy they could be thought of as being anatomical terms."	Technical	Semi-technical
4. Words that have a meaning specific to the field of anatomy and are not likely to be known in general language. They refer to structures and functions of the body. These words have clear restrictions of usage depending on the subject field. Examples are: *thorax, sternum, costal, vertebrae, pectoral, fascia, trachea, mammary, periosteum, hematopoietic, pectoralis, viscera, intervertebral, demifacets, pedicle.*	Technical	Technical

I agreed about **step one**.

With **step two** I have problems with defining the words of position as common. 'Superior' belongs to a set of positional words that includes *dorsal, ventral, lateral, ipsilateral, medial, proximal, distal* etc see "Anatomical terms of location" redirected from http://en.wikipedia.org/wiki/Ipsilateral and as such, while *superior* exists in general English it has a distinctly technical sense in anatomy. Therefore I would call it semi-technical - a word which has a common sense and a technical sense.

Step three is clearly semi-technical. According to Chung and Nation they may be polysemous with common and technical senses.

The confusion continues.

5. **Chung and Nation have ignored the evidence that semi-technical words are a problem for native speakers and non-native speakers**
   I have explained this in the article on terminology. Later in the 2003 article the authors report that "some technical words are common in ordinary English". For which I reply that this is well known and they are often referred to as the semi-technical words. "Applied linguistics has a smaller technical vocabulary and most of that vocabulary (88.4%) is made up of words that are largely familiar to people with no specialist knowledge of the field". This is a highly questionable statement in view of the evidence that semi-technical words are difficult for students.

   Any adequate discussion of technical and common words needs to take into account the difficulty students have with polysemy. The discussion also needs to take into account that highly technical ideas can be expressed in common words, and in doing so the senses of these common words are often modified.

6. **Anatomy is a 'special case'**
   Chung and Nation considered the language of anatomy, which in many ways is a special case, even within medicine. It is certainly full of nouns - the labels for the structures and parts. Anatomy is probably not the most helpful class for analysis. What would be interesting to do is to take the different parts of the body and to see if *within* the domain of anatomy there is a common vocabulary. This method is complicated because within anatomy there are several somewhat overlapping ways of classifying the parts of the body. These include place (object, such as arm), systems (such as the cardiovascular system), and tissue types (such as nervous tissue). Obviously, the 'terms of location' mentioned above would be part of the common vocabulary.

   Within medicine there are other areas that are dominated by nouns. These are:
   a. Chemical nomenclature
   b. Drugs nomenclature

   Other identifiable groups include:
   c. Abbreviations
   d. Acronyms

e. Eponyms. Note, there is a well known trend to reduce these and replace them with a more descriptive term.
f. Names of diseases

In studying anatomy the authors have chosen a subject which is foundational to medicine and the dominance of nouns is not necessarily repeated in other areas. For instance, the Merck Manual of Diagnosis and Therapy http://www.merck.com/pubs/ is a huge comprehensive summary of clinical medicine. This manual is much broader than a textbook of anatomy, and is the kind of reference text likely to be consulted by a wide range of health care professionals.

I am saying then that the research needs to go in both directions. It needs to be more reductionist, and consider smaller parts of anatomy in the attempt to find a useful wordlist for teaching. Research also needs to go in the other direction, avoiding research articles and looking at the broad reference books for the qualified health care professional.

7. Another attempt at classification

In the table below I have accepted the distinction between a word that has more than one sense in the academic world, or is used widely in several disciplines - what Trimble and others refer to as 'semi-technical' (see Terminology, this site, for more discussion and references). I have classified words by sense - following the translators and dictionary makers. In the column 'current usage' I have supplied the terms I would normally used.

Current usage	Common sense	academic sense Trimble's semi-technical	specialised sense unique to one speciality
common words	X		
technical		X	
technical			X
semi-technical	X	X	
technical		X	X
semi-technical	X		X
semi-technical	X	X	X

As a student of mine recently pointed out, in this classification 'technical' is a 'semi-technical' word in the Trimble sense, of having more than one meaning within the academic world. 'Semi-technical' also has the widest range of meanings.

References

Chung TM & Nation P (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language 15(2)* October 2003. Accessed on line http://nflrc.hawaii.edu/rfl/October2003/chung/chung.html on 20 Jan 2007 checked 23.12.09

Chung TM & Nation P (2004) Identifying technical vocabulary. *System 32*: 251-263