Real Science in Clear English

Greg Roos, Murdoch University, Australia
Cathryn Roos, Zayed University, Dubai, UAE

The Asian Conference on Education 2015
Official Conference Proceedings

Abstract

English dominates communication throughout the world of science. For students in countries such as China, India, Asia, the Middle East, and Africa to compete globally in their fields, an increasing number of their tertiary institutes use English – rather than local languages – as the medium of instruction.

However, most available textbooks are written in a manner that assumes total native fluency in English. Studying science is already a challenge, and in a foreign language it is even more so. In our experience, this causes many unnecessary problems in the learning process.

Students need textbooks that are more English language accessible, without reduction of scientific content. This paper outlines our efforts to achieve this. The approach aims to:- eliminate complex sentences, formal vocabulary, and idiomatic expressions; include language features that aid non-native speakers; remove distractions that disrupt readability for students reading in English as a foreign language.

Introduction

It is a well-documented fact that English currently dominates the world of scientific communication. (Gordin, 2015) Throughout the technologically developed world, nearly all serious research reporting is done in English. It is therefore no surprise that the vast majority of popular scientific textbooks emerge from the major English-speaking science communities such as the USA and UK. Since authors in these spheres are primarily native English speakers, the textbooks are often assumed to be the best.

Science students in countries such as China, India, Asia, the Middle East, and Africa wish to compete globally. This usually means competing in English. To help with this, an increasing number of their tertiary institutes now use English – rather than local languages – as the medium of instruction. The vast majority of the books that emerge from the English speaking scientific community continue to assume a native English speaking audience. This creates very real problems in the learning process.

The Non-native Reader of English

What are the challenges that face non-native readers of English? Reading in any non-native language involves two processes: bottom up and top down.(Nuttall, 2006) Bottom up processing involves focusing on letters, words and building up the text. With time and practice, readers develop increasing automaticity in decoding the text. Top down processing refers to what readers bring to the topic, what they draw on of their own knowledge as they go through the text.

Since learning content in a second language will necessarily weigh more heavily on the top down processing, it is critical to make the text as simple and clear as possible in order to support the bottom up processing.

Of course, most non-native English speakers who are enrolled in English-medium tertiary level study require some form of English proficiency qualification, such as TOEFL or IELTS. However, these students are still unlikely to have the same level of automaticity of reading as native speaking English readers. As a result, non-native readers of English often struggle to keep pace with the science course. Their lack of reading fluency limits the focus on actual factual content, and they are in real danger of losing interest in their chosen area of studies.

Current Science Textbooks

Compounding the problem is the fact that scientists are not necessarily good writers. The style and language used in scientific articles tends to restrict readability to a limited audience of like-minded specialists. Greene (Greene, 2013) has recently highlighted this shortcoming, and urges scientists to redress the situation.

In addition to poor text readability, many features of modern textbooks – designed to make the content more engaging, attractive and entertaining – are actually distractions. These distractions can slow down the non-native reader who struggles to sort out the key information from the merely supplementary. Many of the extra features are often meaningless within the cultural setting of the readers.

These extra features and distractions can also be costly. Colorful pages add to the cost of the textbook, as do the inclusion of many of the web-linked activities and various special-interest topics. This is a problem for many non-native readers of English who live in countries where the resources are not available to meet these inflated prices, or to access the many web-based additional materials.

Textbooks for Dummies

Currently there are textbooks on the market that simplify the science, but do not address issues of language accessibility. These books are not appropriate for non-native readers of English in a tertiary setting.

Firstly, these textbooks are aimed at learners who are struggling intellectually with the science concepts whereas non-native readers are challenged by the language, not necessarily the concepts.

Secondly, these textbooks 'for dummies' tend to be much wordier. Also, they include cultural references that may mean very little to the readers. As we can see in the following excerpt from the 'For Dummies' series, the text contains some unnecessarily complex language and uses culturally foreign references to make a point. The italicized parts of the text indicate these issues.

"Carbon *has the capability of forming* four bonds, so molecules that contain carbon can be of varied and *intricate* designs. Also, carbon *bonds represent the perfect trade-off* between stability and reactivity - carbon bonds are neither too strong nor too weak. Instead, they epitomize what chemists *refer to affectionately as the Goldilocks principle* - carbon bonds are *neither "too hot" nor "too cold," but are "just right."* If these bonds were too strong, carbon would be unreactive and useless to organisms; if they were too weak, they would be unstable and would be just as worthless."(Winter, 2014)

Real Science, Clear English

Since any reduction of the scientific content is not usually an option, the only alternative is to improve the readability.

To address these issues, we developed the concept and principles of Real Science Clear English. (Roos & Roos, 2015b). Many of these principles are drawn from guidelines for good writing in general; (Greene, 2013) some of these principles are drawn from the Plain English concept, but with some important differences. ("Plain English Campaign," 2015)

The 'Plain English' concept, is largely aimed at native English speakers, usually in the areas of public information and business. Clear English, on the other hand, is specifically aimed at non-native readers of English and remains academic in nature.

Although both approaches embrace clarity and brevity, material for non-native readers of English needs to be written in clear English without compromising on academic rigor. The Clear English concept for textbooks means: simple, uncluttered layout with no unnecessary distractions, such as sidebars or italic fonts; the simplest sentence constructions possible, even if monotonous to the native reader; a minimum of different vocabulary items, even if the same words for the same things are used over and over; no idiomatic expressions or heavily cultural references.

Principles of Clear English

A few of the principles of 'Clear English' are illustrated in the following excerpts. We begin with a 'before' paragraph from the original introductory organic chemistry textbook (REF) that we transformed into a language-accessible version of the same textbook. (Roos & Roos, 2015a). We finish with the 'after' paragraph to show how the language can be made clearer.

The original text: 'Before'

Before beginning any discussion of the compounds based on carbon, it is perhaps appropriate to put the element into perspective in terms of its natural abundance in a system such as the human body. (Table 1.1) Abundance alone does not dictate the *importance of the elements of life. Iron is part of haemoglobin which carries oxygen in the blood. Iodine is vital to smooth operation of the thyroid. Cobalt is contained in vitamin B_{12}. Zinc, copper, and manganese are present in various enzymes. In each of these examples there are vast numbers of carbon, hydrogen, oxygen, and nitrogen atoms for every metal atom. Nevertheless, without the metal, the compounds could not serve their biological functions.*

Principle 1: Remove unnecessary meta-data

Before beginning any discussion of the compounds based on carbon, it is perhaps appropriate to put the element into perspective in terms of its natural abundance in a system such as the human body. (Table 1.1) Abundance alone..

~~Before beginning any discussion of the compounds based on carbon, it is perhaps appropriate to put the element into perspective in terms of its natural abundance in a system such as the human body~~. (Table 1.1) Abundance alone..

Principle 2: Use appropriate simple clear visuals

…. in a system such as the human body. (Table 1.1) Abundance alone does not dictate the importance of the elements of life.

Table 1.1 shows examples of trace elements in the human body.

Principle 3: Use concrete rather than abstract nouns

Abundance alone does not dictate the importance of the elements of life. Iron is part of haemoglobin which carries oxygen in the blood. Iodine is vital to smooth operation of the thyroid.

We have only a small amount of iron in the body. However, iron is needed to carry oxygen in the blood. Iodine is needed for the thyroid to work.

Principle 4: Use simple, high-frequency vocabulary

Iron is part of haemoglobin which carries oxygen in the blood. Iodine is vital to smooth operation of the thyroid.

Iron is needed to carry oxygen in the blood. Iodine is needed for the thyroid to work properly.

Other principles incorporated into the new textbook include:

- Highlight key terms/phrases for a glossary
- Keep sentence length to maximum of 20 words
- Avoid lengthy noun clauses
- Use simplest sentence type possible and limit the range of linking words:
  - Simple
  - Compound with limited linking words - and, but, or
  - Complex with limited linking words – because, although, which

The new text: 'After'

The amount of each element of life does not tell us how important it is. Table 1.1 shows examples of trace elements in the human body. We have only a small amount of iron in the body. However, iron is needed to carry oxygen in the blood. Iodine is needed for the thyroid to work properly. Cobalt is part of vitamin B_{12}. Zinc, copper, and manganese are in various enzymes that the body needs. In each of these examples there are many carbon, hydrogen, oxygen, and nitrogen atoms for each metal atom. However, without the trace element metals, it is impossible for these compounds to carry out their biological functions.

Tools for Assessing Language Accessibility

A number of useful web-based tools exist to evaluate the difficulty and level of readability of text. We used several of these to evaluate both the above 'before and after' sample text, as well as the full pilot textbook. Snapshots of evaluations are shown in Figures 1-3 below, and the key indices are collected in Table 1.

Tool 1: Coh-Metrix

This tool gives various indices of the linguistic and discourse complexity of a text. There are 106 indices, but we focus on the one for ESL readability (the last one). The higher the score, the easier to read.

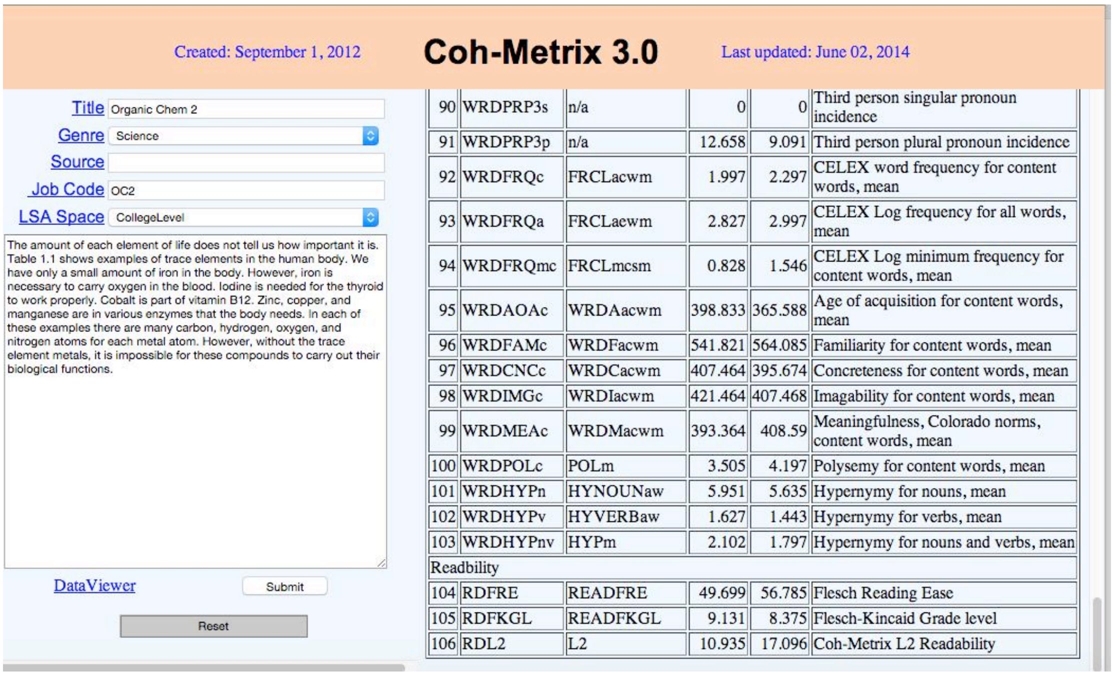

Figure 1 Coh-Metrix sample output(McNamara, Louwerse, Cai, & Graesser, 2013)

Tool 2: Average Grade Level

This tool is based on American school grades. The lower the score, the easier to read.

Grade Levels

A grade level (based on the USA education system) is equivalent to the number of years of education a person has had. A score of around 10-12 is roughly the reading level on completion of high school. Text to be read by the general public should aim for a grade level of around 8.

Readability Formula	Grade
Flesch-Kincaid Grade Level	8
Gunning-Fog Score	11.6
Coleman-Liau Index	10.2
SMOG Index	8.3
Automated Readability Index	5
Average Grade Level	**8.6**

Figure 2 Composite readability grade level sample evaluation("Readability," 2015)
Tool 3: The Vocabulary Profiler
This tool breaks down the text into levels of vocabulary frequency in the language at large. There are different vocabulary lists that can be used. This one is from the new general service and new academic word lists (NGSL/NAWL).

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token %
K-1 Words :	25 (44.64)	26 (43.33)	44 (55.70)	55.70
K-2 Words :	7 (12.50)	7 (11.67)	8 (10.13)	65.83
K-3 Words :	11 (19.64)	12 (20.00)	12 (15.19)	81.02
K-4 Words :	5 (8.93)	5 (8.33)	6 (7.59)	88.61
K-5 Words :	2 (3.57)	2 (3.33)	2 (2.53)	91.14
K-6 Words :				
K-7 Words :	1 (1.79)	1 (1.67)	1 (1.27)	92.41
K-8 Words :	2 (3.57)	2 (3.33)	2 (2.53)	94.94
K-9 Words :				
K-10 Words :	2 (3.57)	2 (3.33)	2 (2.53)	97.47
K-11 Words :	1 (1.79)	1 (1.67)	1 (1.27)	98.74
K-12 Words :				

Figure 3 Sample evaluation of common English words(Cobb, 2015)

Table 1 Key evaluation indices of 'Before and After' sample text

Readability Tools	Before	After
Coh-metrix readability (RDL2-L2)	10.935	17.076[a]
Average grade level (based on 5 readability indices related to American school grade levels)	10.6	8.6[b]
1000-2000 most frequent words (specifically, the New General Service List and the New Academic Word List	66%	80%[a]

[a] The higher the value, the better the readability
[b] The lower the value, the easier the readability

As can be seen from Table 1, all the indices change significantly in the direction of improved readability. These same evaluations were carried out on the entire pilot publication(Roos & Roos, 2015a) and revealed that, except for specific technical content, only about 1250 different English words were used. Of these, more than 98% are drawn from the 1000-2000 most common English words along with the 350 strong Academic Word List.

Conclusion

Late secondary and entry level tertiary non-native English readers who study science will benefit greatly from textbooks specifically written in Clear English. These students, as they progress through the tertiary system, will have a greater reading confidence and will be able to progress more easily to the mainstream textbooks required in their senior years.

We need to convince both authors and publishers to strive toward producing more appropriate introductory science textbooks in Clear English. If this does not happen, we risk losing many good students due to the frustration of reading unnecessarily complex writing. An alternative is to move away from the monoglot predominance of English in science and a return to some new polyglot of languages. This less preferable outcome will surely result in overall reduced global communication in the important areas of science and technology.

References

Cobb, T. (2015). Compleat Lexical Tutor (Version v.8). Retrieved from http://www.lextutor.ca

Gordin, M. D. (2015). *Scientific Babel*. Chicago: University of Chicago Press.

Greene, A. E. (2013). *Writing Science in Plain English*. Chicago: The University of Chicago Press.

McNamara, D.S., Louwerse, M.M., Cai, Z., & Graesser, A. (2013). Coh-Metrix (Version 3.0). Retrieved from http://www.cohmetrix.com

Nuttall, C. E. (2006). *Teaching Reading Skills in a Foreign Language* (3rd ed.). Oxford: MacMillan.

Plain English Campaign. (2015). from https://www.plainenglish.co.uk/

Readability. (2015). Retrieved from https://readability-score.com

Roos, G. H. P., & Roos, C. L. (2015a). *Organic Chemistry Concepts: An EFL Approach*. Boston: Academic Press.

Roos, G. H. P., & Roos, C. L. (2015b). RealScienceClearEnglish, from http://www.realscienceclearenglish.com

Winter, A. (2014). *Organic Chemistry I for Dummies, p.14* (2nd ed.). New Jersey: John Wiley & Sons.

Contact email:
g.roos@murdoch.edu.au
c.roos@zayed.ae
realscienceclearenglish@gmail.com