Natural Language: Application of its analysis to the improvement of student authoring skills

Rick Duley
North Perth, Western Australia
rickduley@gmail.com

Abstract

One of the greatest assets a tertiary education can bestow on an intending Engineer is the ability to readily and efficiently produce and assimilate technical communications. It has been suggested that such work will occupy 50% of an Engineer's professional time.

This document presents the concept that analysis of the language used in student-authored documentation will provide information on which pedagogy of technical authorship may be based. It presents the importance of the case and looks at the data involved and some of the means of analysing it.

1 Introduction

Rudolf Flesch, creator of the Flesch Reading Ease Scale, started Chapter Two of his book *"How to Write Plain English"* as follows:

> *"I've come across quite a few people who can write a booklet explaining an insurance policy because they have the knack. But I take it you're not blessed with that knack. You belong to the 98 or 99 percent of mankind who must learn Plain English the hard way."* (Flesch, [web page])

Here Flesch highlighted two major truths about writing: (1) that the vast majority of people are not good at it and (2) that writing is a skill that can be learnt. True, all graduates of an English Literature programme (or an Engineering programme) will not be literary giants, but there must be hope that they will be at least competent — that is they will be able to readily and efficiently read and write the types of documents they will encounter or will have to produce throughout their careers. After all, quality of software development documentation impacts directly on the quality of the software produced. Recent research found that 56% of software bugs originated in the Requirements phase and 27% in the Design phase — both phases being documentation intensive. Only seven percent of bugs originated in the Coding phase. (See Figure 1, (Mogyrodi, 2003).)

In fact there are two conflicting schools of thought about the ability to write well: firstly, that writing is an art-form

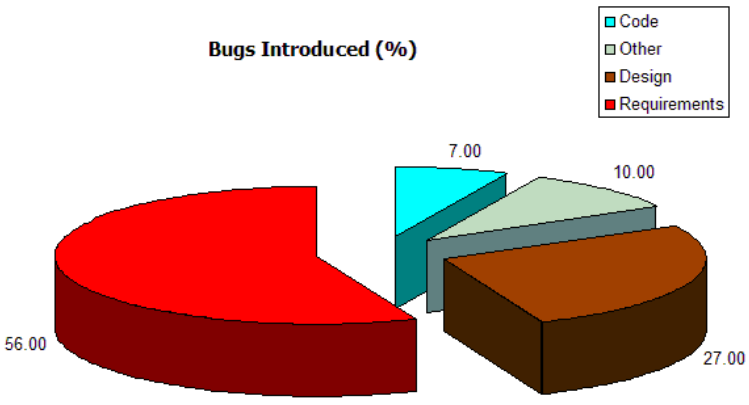

Figure 1: Origins of Software Bugs

and that writing talent is a gift, something that cannot be taught; secondly that it is a teachable craft (*'wordsmithing'*) (Levine, Pesante, & Dunkle, 1991, p. iv). I hold to the belief that each affirms a truth; those with the talent will usually excel over the rest no matter how much learning either group undertakes but this should not restrain us from seeking to advance the abilities of those without the natural talent.

Writing, however, cannot be reduced to the observance of a set of rules:

> *"There are no algorithms for writing. Writing is rhetorical, requiring the writer to take into consideration many factors that change from one situation to the next."* (Levine et al., 1991, p.iii)

Even less can it be reduced to a process of modification to match a required formulaic outcome; as we shall see the best of readability formulas can only serve as a guide. Levine et al. go on to state the aim of their document as being to provide strategies or sets of heuristics to aid the reader to become a more effective writer. Such aims underpin this document and its support work.

1.1 Engineers and Communication

There is little doubt that communication skills are vital professional attributes in Engineering. For examples:

- When employers were asked, in 2003, to rate prospective employee qualities and skills on a five-point scale (1 being *'not important'* and 5 being *'extremely important'*) Communication Skills rated an average of 4·7 (IEEE/ACM Joint Task Force on Computing Curricula, 2003, p.15);
- Huckin pointed out that engineers spend 50% of their time communicating (Huckin & Olsen, 1991, p.17);
- CCSE lists written and oral communication as integral components of professional practice (IEEE/ACM Joint Task Force on Computing Curricula, 2004 Ch2.1[10], p.10);
- CC2001 cites *'written and oral communication'* as facets of professional practice which must be included in an undergraduate curriculum (IEEE-CS/ACM Joint Curriculum Task Force, 2001, section 4:10, p.13).

This valuation of communication skills is not confined to the engineering realm either: in Australia the Mayer Report listed as a Key Competency,

> *"The capacity to communicate effectively with others using a range of spoken, written, graphic and other non-verbal means of expression."* (Mayer, 1992, p.3);

In the UK communication is listed as one of three *basic* skills comprising a national Key Skills Qualification (Australian National Training Authority, 2003, p.6); and Teichler points out that,

> *"... the most outspoken voices claim that graduates should...cultivate social and communicative skills..."* (Teichler, 1999, p.298).

While the term *'Communication Skills'* includes writing skill, it encompasses

much more than that. Huckin and Olsen, in listing 38 subjects most needed for engineering careers in industry, named (with ranking) the following characteristics — all of which fall within the compass of *Communication Skills*:

- Management practices (1);
- Technical writing (2);
- Public speaking (4);
- Working with individuals (6);
- Working with groups (7);
- Speed reading (8);
- Talking with people (9) (Huckin & Olsen, 1991, p.5).

Snoke and Underwood listed as the top seven desired generic competencies (rated on a seven-point Likert scale where 1 = extremely unimportant and 7 = extremely important):

- Work as part of a team in a productive and cooperative manner (6·30);
- Be able to participate in continued learning and intellectual development and develop critical, reflective and creative thinking (6·28);
- Retrieve, evaluate and use relevant information (6·22);
- Oral communication skills (6·17);
- Define problems in a systematic way (6·11);
- Written communication skills (6·09);
- Interpersonal skills (6·02). (Snoke & Underwood, 1999, p.820)

Six of those seven are communication intensive!

1.2 Teaching Communication to Engineers

Judging by this it is safe to assume that a considerable amount of Huckin's *"50%"* of an Engineer's time will involve Communication. On the basis of that assumption there is a solid claim for the application of much more teaching time to the preparation of technical documents[1] than is currently the case.

> *"The premise is that the development of written communication skills is too important to be left solely to one or two specific courses. Instead, students need to practice and reinforce these skills in the context of many different courses throughout the student's educational experience. The English Department is, in effect, saying that they can provide the foundations, but we must all share in this important task."* (Gersting and Young, 2001, p.18)

> *"...effective writing is not something that can be covered once and mastered... we stress the need to integrate writing into the curriculum, making it part of each course that the students take."* (Levine et al., 1991, p.iii)

However, the fact is that writing skills remain a *poor cousin* in the curriculum, paid lip service but regularly overlooked in resource allocation.[2]

[1] In this context, the term 'technical documents' should be taken to mean any documentation appropriate to the discipline of Engineering. This would include requirements, user or maintenance documentation on the one hand and the preparation and delivery of presentations or journal articles on the other.

[2] CUT has taught a unit entitled "Developing written and oral communication skills in Engineering students" as a 'generic first year subject' since the early 1990s (Australian Technology Network, [web page]) and RMIT has a first year unit it "Written and Oral Communication" (RMIT, [web page]).

"Why do technical professionals often communicate so badly? One reason is poor training. In trying to keep up with rapid advances in technical knowledge, technical schools are often forced to make sacrifices and sometimes it is the teaching of communication skills that gets shortchanged — despite repeated complaints from industry. " (Huckin & Olsen, 1991, p.13-14)

Student writing capabilities are often seen as being low when students leave High School:

"Most students are poor writers. July 2003 NAEP[3] test scores show that fewer that one in three pf the nation's fourth, eighth and 12th-graders are proficient in writing — that is, capable of composing organised, coherent prose in clear language with correct spelling and grammar. Only 24 percent of high school seniors achieved that goal. … In a 2002 public agenda survey, more than 70 percent of employers who hire recent high school graduates and college professors who teach freshmen and sophomores rated public high school graduates 'fair' and 'poor' on writing." (Hurwitz & Hurwitz, 2004)

When students leave university, matters do not seem to have improved much:

"… employers use academic results as the initial basis for selection, but after that they select graduates on the basis of their communication skills … These skills were perceived to be generally lacking in new graduates…" (Barthel, [web page])

Writing skill, as in the ability to make what is written readable, has long been seen as rare:

"Let's face it. Most of our documentation is dull. We explain facts, when people want recommendations. We provide how and what, but not why or who. Our writing is not musical or dramatic. The feeling associated with lists and forms is boredom. We try to make our writing neutral, and end up taking the personality out of it. The addition of the second person singular does not put people into our writing. The nameless 'you' only invokes the shapeless 'one'. … We try so hard to be objective that we lose all the flavour. … the process of learning should be fun as well as convincing." (Hallgren, 1986, p.109)

Without effort on the part of Universities there is no reason to believe that student writing capabilities will improve before graduation or that they will they will blossom into literary splendour on employment.

1.3 Engineers' Writing in the Learning Process

Engineering graduates must understand their subject. It is not enough that they can cite and manipulate formulae or quote physical laws, they must understand their subject in the way a doctor understands physiology, being able to transfer technical data into a practical situation:

"Educational institutions, of course, also tend to neglect meaningful learning; medical schools, for example, suffer from a reputation of inculcating basic science into student's heads, then expecting them to successfully transfer that knowledge into clinical settings later." (Wilson, Jonassen, & Cole, 1993)

It is said that the principle applied in medical colleges is, *'Watch one; Do one; Teach one.'* As far as students are concerned writing may fill the place of the '*Teach*' in the learning process:

"If students are to make knowledge their own, they must struggle with the details, wrestle with the facts, and rework raw information and dimly understood concepts into language they can communicate to someone

[3] (US) National Assessment of Educational Progress

else. In short, if students are to learn, they must write." (Hurwitz & Hurwitz, 2004)

This document looks at a way to provide an empirical evaluation of the results of teaching writing skills, and a means to track trends in those results, as an advance towards graduating engineers with substantial skill in written communication.

2 Writing

Self-evidently, the most brilliant and creative of engineers will remain ineffective if they are unable to pass their vision to others. Given that most engineers work in situations where they must communicate with clients, sponsors, politicians and other vested interests, engineers must not only be able to communicate with their peers but also with people who have little or no understanding of the engineer's particular field. Many methods exist to make this possible, but diagrammatic and writing competence must figure high on any list. This raises the question of how writing competence might be judged:

"...those who assess writing must ask themselves what observable traits or behaviours constitute good writing." (Deacon, [web page])

2.1 Good Writing

Researchers have identified many signs of good writing and a brief list includes:

- Average number of words in a sentence;
- Correct emphasis;
- Correct grammar;
- Correct punctuation;
- Enjoyability;
- Keeping to the subject;
- Logical reasoning;
- Minimal jargon;
- Number of personal pronouns;
- Number of prepositional phrases;
- Omission of needless words;
- Organisation and structure of material in an orderly and logical way;
- Percentage of different words;
- Range of vocabulary;
- Selection of target audience;
- Use of illustrations, tables and lists;
- Use of language that is simple, direct, economic and familiar;
- Use of sentence structures that are evident and unambiguous;
- Voice — active, imperative.

Note: the above list was garnered from a few documents: (Everything You Ever Wanted To Know About Readability Tests But Were Afraid to Ask., [web page]); (Huckin & Olsen, 1991); (How to Write in Plain English, [web page]) and (Hargis, 2000). It is unlikely that there is such a thing as a complete list!

Attempts to define Good Writing in terms of empirically identifiable values such as these appear, unfortunately, to founder on the rhetorical aspect of writing referred to in section **Error! Reference source not found.**. After all writing developed long after oral

communication and remains, arguably, a fixed version of its essentially fluid predecessor — as Deacon cites:

> *"...the greatest attraction of writing is its subjectiveness or anarchic amorphousness."* (Deacon, [web page])

Furthermore, English is a living language in which vocabulary, syntax and semantics are in constant flux. Time-honoured, rigid rulings become increasingly inappropriate:

> *"You can start a sentence with and, but, because, so, or however. You can split infinitives... You can end a sentence with a preposition... And you can use the same word twice in a sentence if you can't find a better word."* (How to Write in Plain English, [web page])

How then is one to identify Good Writing if one can no longer look for — and, supposedly enumerate and analyse — characteristics like those listed above? How are such indeterminables as the effectiveness of illustrations factored in?

Reductionist tendencies, which would have the definition of Good Writing reduced to such numerical analysis, have in the past become objects of scorn. Hargis described these tendencies as being similar to *'putting a cloud in a bottle'* (Hargis, 2000, p.123). If it is true, then, that Good Writing can only be detected by subjective appreciation of what is written, of how it is written and in the context in which it is written, is it then possible to make an objective assessment of Goodness in Writing? Andrew Forward provides a key to a final assessment:

> *"Based on the assumption that documentation is communication, the goal of a particular software document is to convey information."* (Forward, 2002, section 2.1.1, p.7)

This might be a statement of the self-evident but in the end, when it comes to judging a document, little else matters. Writing works if it conveys information. If it works, it's good!

This document is, therefore, based on the definition of good writing as *'that which conveys the necessary information'*. Good documentation, I reassert, contains information that is comprehended by the reader. Good documentation must be readable. Many factors govern how readability is achieved.

3 Readability

Readability is often confused with legibility or even comprehensibility. While legibility deals with whether or not your handwriting can be deciphered and comprehensibility with whether or not you have anything sensible to say, readability measures how comfortably or easily your work may be read (Gregory, [web page]). This comfort depends on a variety of factors such as sentence length:

> *"When you read a passage, your eyes and mind focus on successive points on a page. Each time this happens, you form a tentative judgement of what the words mean up to that point. Only when you get to a major punctuation mark — a period, a colon, a paragraph break — does your mind stop for a split second, sum up what has been taken in so far, and arrive at a final meaning of the sentence or paragraph. The longer the sentence, the more ideas your mind has to hold in suspense until its final decision on what all the words mean together. Longer sentences are more likely to be complex — more subordinate clauses,*

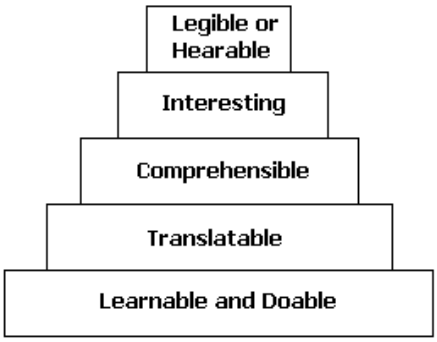

Figure 2: Levels of Readability according to Hargis

more prepositional phrases and so on. That means more mental work for the reader." (Flesch, [web page])

Considerable research has been undertaken into the concept of defining Goodness in Writing in terms of reducing that mental work, *'Readability'*. Computations designed to assess readability are usually weighted linear formulae based on readily measurable factors:

"Today readability formulas are usually based on one semantic factor (the difficulty of words) and one syntactic factor (the difficulty of sentences). Studies have confirmed that inclusion of other factors in the formula contributes more work than it improves the results." (Everything You Ever Wanted To Know About Readability Tests But Were Afraid to Ask., [web page])

Unfortunately it has also become obvious that:

"...the underlying assumption of readability formulas — that any text for any reader for any purpose can be measured with the same formula — does not mesh with our current understanding of how people read and understand."[4]

This is obviously a complex field and for this reason I digress to give a brief overview of some of the concepts of Readability.

3.1 Levels of Readability

Hargis proposed an hierarchy of five readability levels (Figure 2) ranging from *'Legible or Hearable'* to *'Learnable and Doable'* (Hargis, 2000, p.123). She made the point that extra reader-effort is required to maintain focus on uninteresting material then went on to comment that:

"The people who need [computer] *information have no choice but to read what's offered, and any techniques to make the information interesting may even be resented as an affront to the user's professionalism."* (Hargis, 2000, p.123)

Two issues arise from this observation:

- People who *do* read technical documentation do so because they need the information it is assumed to conceal. There is, in that sense, no question of the degree to which the information may be made to *appear* interesting — the information, *de facto*, ***is*** interesting. If it were not so, the documentation would not be being read. Given that there is a high level of interest intrinsic to technical documentation, the hierarchy might be better presented as in Figure 3.
- There may be a certain degree of snobbery in the idea that technical or scientific writing should be hard to read. I vehemently disagree that it should. I believe that it is time that that hoary old excuse for communicative incapability was shot with a silver bullet and gently laid to rest impaled with a wooden cross!

[4] Redish, J.C. and Selzer, J. (1985) The place of Readability Formulas in Technical Communication. Technical Communication, 32(4):46-52 cited in (Klare, 2000, p.153).

There is no excuse for stodgy, awkward prose that, by being so, pretends to be scientific! Furthermore, that any supposed professional (either as writer or reader) should take pride in the concealment of professional truth under legalese, jargon or plain, old-fashioned bad writing is contradictory to professionalism.

> *"Perhaps professionals who routinely write at college or graduate school reading level are frightened by the recommendation that they write at 8th grade reading level, so they attack plain language efforts that they cannot personally achieve."* {Hochhauser 1999 #1332 /ft "p.25""}

- Authors who hold this belief might take stock of the fact that they are authors, whose service is to communicate, not readers who need the information (and have paid in the hope of getting it).

> *"*[Q] *But I have a good education, why do I have to write below it?*
>
> *"*[A] *Because you are the writer; not the reader. We don't write to serve ourselves. We write to serve our readers; all of them, not just the most skilled among them. With your education, you will find it easier to write well and still be easily read."* (Readability: Frequently Asked Questions, [web page])

Looked at another way, those who have so little certainty of their own grasp of professional knowledge as to feel the need to shield that knowledge from the understanding of outsiders and novices can only cheapen the reputation of their profession. Perhaps these unfortunates should note, as did Hargis, the kudos accorded some retail computer books on the grounds of their helpfulness and understandability (Hargis, 2000, p.124). They might remember that the founders of International Data Group, publishers of the *"... for Dummies"* books, were able to donate US$350m to MIT to create an Institute for Brain Research (Klare, 2000, p.149).

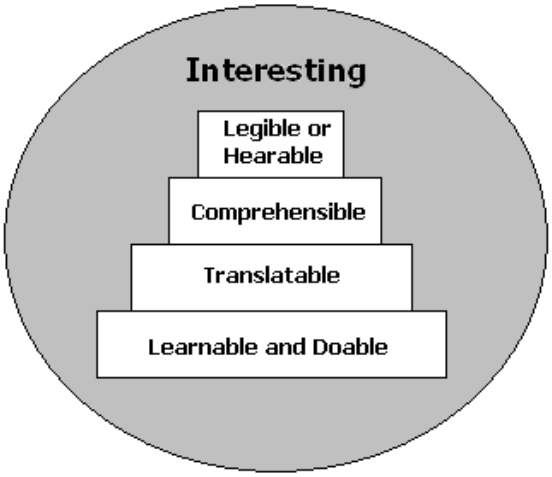

Figure 3: Levels of Readability Assuming Sufficient Level of Interest

3.1.1 A Question of Readability

In the case of technical documentation, it is safe to assume that a sufficiently high level of interest in the subject matter pre-exists the reading. Removing *'Interesting'* as a quality of readability from the taxonomy, and regarding it as naturally pervasive in this context, leaves us with just four levels of readability to consider. These four levels (Figure 3) follow a logical pattern:

- That which is not legible is not comprehensible;

- That which is not comprehensible is not translatable;
- That which is not translatable is neither learnable nor doable.

3.1.1.1 Readability as Legibility

Whether we are talking about legibility or audibility we are talking about the ability of the communication medium to provide sensations which can be conveyed to the brain as data for conversion to information. Although there are exceptions (e.g. the archetypal doctor's prescription) communication media are created in accordance with certain conventions. These conventions make the innate content of the medium readily available to the educated eye or ear. As far as technical documentation is concerned, flowing, copper-plate handwriting is rare these days but that matters little. In its place, capability with a word processor becomes fundamental. That of itself, however, is not a complete solution; something as simple as the selection of the font (typeface and size) becomes highly influential. As the Plain Language at Work Newsletter put it:

> *"During the past thirty years, there has been a great deal of interest and research in the graphic aspects of writing that appeal to the right side of the brain. They include typography, editorial design, layout, symmetry, and the generous use of illustrations, colour, white space, graphs, and bulleted lists."*

However, the newsletter does go on to point out:

> *"While good design is helpful in attracting a reader to a text, it does not guarantee understanding of the text. For that, you need the right words lined up in the right order to match the skill of the reader. If the words are too long or if the sentences are too complicated, you will lose the reader quicker than it takes to turn a page."* (Plain Language At Work Newsletter, 2004)

So, familiarity and facility with layout, typography, consistency, colour and coordination stand out as important although not the *'be-all and end-all'*. However, given a grounding in these, and high end-user status with a word processor or desktop publishing package, the professional should be able to present documentation which is pleasant to view, ergonomically easy to read and provides the intuitive navigational insights into the structure of the document.

3.1.1.2 Readability as Comprehensibility

Four decades ago research did not show clear correlation between readability and comprehension (Hargis, 2000, p.124) and readability could be defined as the ease with which material was read (although not necessarily understood) (Klare, 1963). By the 1980s readability and comprehensibility were synonymous according to (Klare, 1988) yet there remains a transfer between legibility (audibility) and comprehensibility involving more than just readability. It involves two aspects, neither of which is addressable in terms of document construction.

> *"Although the* [U.S. Federal Drug Authority] *requires that the information in the consent form be presented in a language that is understandable to the subject, the ability of the subject to understand the consent language is not the same as the ability to understand the consent concepts."* {Hochhauser 1997 #1333 /ft ", p.8"}

They involve the use of the (speaker's/writer's) language in a more metaphysical sense:

- **Language Transfer:** Hargis suggests that native users of English are outnumbered 2:1 by those who use English as a second language (ESL) (Hargis, 2000, p.124). That still leaves twice the number of people who do not speak English at all as there are native English speakers, and for these people (many of whom will be seeking computer literacy) documentation will require translation. For either group of non-native English speakers, reduction in levels of complexity in technical documentation and removal of ambiguity from that documentation assume considerable importance. ESL *writers*, writing for native English speakers, have a compounded problem in that the readability of their document will be

Example 1: Revision for Comprehensibility

The following text is taken from an expensive four-colour mailer advertising the programs of Hoag Hospital in Newport Beach, California. It is written at post-graduate level:

"Maintaining balance and flexibility are two key elements that help patients feel better and move more freely. Both are generously addressed through Hoag's comprehensive physical and occupational therapy programs, which offer targeted therapies in gait and balance treatment, as well as fall-prevention training for patients with difficulty walking. Hoag also offers a specialised program for combating the swallowing problems that often accompany Parkinson's disease, and an innovative speech therapy program aimed at improving voice projection and diction."

The following re-write contains the same information, but is written at the 7th-grade level:

"Balance and flexibility help people feel better and move freely. If you know someone who has problems getting around, contact Hoag Hospital. We have good programs that can help people walk and move better, without fear of falling. Parkinson's Disease can also cause problems with speech and swallowing. We can help people with those problems too."

This revision can be read by 75% of the adult population, causing a significant increase in readership. (Plain Language At Work Newsletter, 2004)

subjectively judged by native English speakers on the basis of their fluency. Unfortunately, fluency is one of those things that appears to have eluded evaluation by computer programs.

- **Content Transfer:** Consider the hypothetical cases where (1) an Electrical Engineer describes a technological advance or (2) a Computer Scientist describes a new algorithm. Each must consider the audience. Word selection must vary to suit, on the one hand, the specialist audience fully familiar with the subject and, on the other, the novice audience struggling to keep up with the basics. What, however, of a Conference Audience, or the readership of a book, almost certainly made up of a mixture of both types? This is an aspect of *translation* dealt with in greater detail in Section 3.2.

Both of these issues are vital to making the information contained in written material comprehensible, but comprehending information doesn't guarantee personalisation or ownership of that information.

3.1.1.3 Readability as Translatability

Technical documentation, as the term should be taken in this document, is concerned with information which readers will use to do something. For this reason, the information must become part of the reader's self – it must become something that the reader knows in the sense that *knowing* is different from merely *believing something to be factual*. This is the translation from the *'know what'* to the *'know how'*.

> *"Many methods of didactic education*[5] *assume a separation between knowing and doing, treating knowledge as an integral, self-sufficient substance, theoretically independent of the situations in which it is learned and used. The primary concern of schools*[6] *often seems to be the transfer of this substance ..."* (Brown, Collins, & Duguid, 1989)

In technical documentation, readability involves the presence of the means of ownership transfer; technical documents must allow the reader to take possession of the knowledge they contain. Failure results in the possession of *inert knowledge*:

> *"Sir Alfred North Whitehead (1929) coined the term* 'inert knowledge' *for the kind of knowledge schools typically teach. Students often fail to use knowledge gained in one setting (schools) in another key setting (on the job). Thus their knowledge is inert and is of no use to them when they need it. ... This is called the transfer problem. ..."* (Wilson et al., 1993)

This *inert* characteristic of knowledge is fundamental to the *'banking'* concept of education which holds that the teacher (or document) holds knowledge and the function of the student (or reader) is to collect and catalogue, to *file away* that knowledge. Seeing knowledge as inert does not acknowledge the phase of translation from the comprehensible to the doable. As Paulo Friere put it:

> *"... in the last analysis, it is men themselves who are* [thus] *filed away through lack of creativity, ... The more* [people] *work at storing the* [knowledge] *entrusted to them, the*

[5] Reading technical documents is seen here as a form of self-education.
[6] And technical documents?

less they develop the critical consciousness which would result from their intervention in the world as transformers of that world." (Friere, 1972, pp.46,47)

Engineers are transformers of their environment, not mere dwellers in it. Therefore, technical documentation must support the transmission of knowledge for the purpose of action.

3.1.1.4 Readability as Doability

Readers of technical documentation generally require near-100% comprehension as a basis of performing some task (a) for themselves and (b) immediately. While High School textbooks may be studied for the purpose of passing an examination, the concept of examination takes on an entirely different meaning when, in the post-school world, *'passing'* means keeping one's job (or even, perhaps, one's life).

"It's not enough that the reader be able to understand, learn for the sake of learning, or read faster. Such purposes in the mind of the writer are more appropriate for a student audience than for an audience of users of computer products and computer documentation.

"For technical communicators working on computer documentation, our purpose is to support the user's tasks in using the product. The ultimate test is whether the information or assistance satisfies the user's needs, preferably before they get frustrated with the product." (Hargis, 2000, p.126)

3.2 Readership Characteristics

In any event, as Klare pointed out forty years ago:

"If it is not readable to an intended reader it is not readable, no matter how good a formula score it may receive." (Klare, 1963, p.11)

Four decades later he wrote:

"Studies already make clear that certain reader characteristics relate to the comprehension of documents. A high level of reading skill, when present, can provide a high level of reader comprehension. A high level of motivation can be shown to override some comprehension difficulties, and a high degree of background knowledge can make even difficult documents comprehensible. <u>But documentors cannot always count on high levels of reader skill, motivation and background</u>[7]." (Klare, 2000, p.165)

Any author or speaker does well to take into account the characteristics of the intended audience, however, in the case of the readers of technical documentation, many of the common variables in the group are removed. Johnson points out that readability involves a third party matching reader and text (Johnson, 1998, p.1) but in the case of technical documentation the issue is of subtly different[8]. In this case either the readers <u>make the choice</u> or <u>there is no choice</u>. Interest, attention and motivation exist on the basis of the need to know as mentioned in Section 3.1. Readable text simply removes an obstacle to the satisfaction of that need. This does not, however, mean that the readership is homogeneous to the point of identity:

"...users of some personal computer documents range from the nerd to the scared." (Klare, 2000, p.150)

[7] The emphasis is mine - rd

[8] Johnson was seeking textbooks already written suited to specified age groups, while the issue here is writing for a largely unspecified readership.

3.2.1 Level of Base Knowledge

Almost certainly, the greatest variance between readers will be their level of understanding of the topic before they read the document. Technical documenters cannot assume any standard of reader foreknowledge except the lowest in the expected audience group, which means running the risk of alienating the more pre-accomplished reader with reiteration of fundamentals. Many technical documents avoid this possible alienation by advising power-users to *'skip chapters two and three'* or some similar device.

3.2.2 Demographic Aspects

Globalisation of the software industry means that technical documents must travel well. Language, culture, religion, environment etc. will all colour the reader's interpretation of a document, yet is has long been known that ideas can travel well if those issues are understood and well handled. Hofstede showed that, even with widely varying cultural factors, management principles and practices could be readily transferred from one setting to another (Hofstede, 1984). For example, much was being made at that time of the influence of Japanese culture on the success of the Kan Ban manufacturing system in Japan. Common commentary was that, *'it can't work here'*. However, in support of Hofstede's conclusions, it is readily demonstrable that Kan Ban was transported successfully to a host of nations. It is important to realise that Hofstede did not suggest that the cultural differences were of no importance, simply concluding that:

> *"Effectiveness within a given culture, and judged according to the values of that culture, asks for management skills adapted to the local culture."* (Hofstede, 1984, p.98)

I would argue not only that technical ideas are more readily transportable than possibly-emotive management theory but also that demographic awareness of the target audience will facilitate the interchange.

3.2.3 Motivation

Enthusiasm for the topic (as distinct for the need for the knowledge referred to on page 7) influences the effort a reader may be prepared to put into deciphering difficult text. Consider:

> *"Readability indexes are helpful, but they do not consider the best motivator: interest in the material. People will read at a much higher level the things they find interesting."* (Wheeler, [web page])

or

> *" A young electronics enthusiast may read and persevere with a complex electronics magazine, but quickly abandon the simplest history book. This internal motivation is very powerful, but not easily modified by a teacher."* (Johnson, 1998, p.1)

Interest is a powerful motivator — however, as we have seen in Section 3.1 it may be assumed to be omnipresent in the case of technical documentation.

4 Evaluating Readability

Traditional, manual, evaluation of writing involves the evaluation of a range of indicators that are, in

themselves, insubstantial and which prove extremely difficult to define sufficiently to enable automation. Consider the criteria established by the Illinois State Department of Education in the US for profiling a student's command of clear writing:

- **Focus:** Is the main idea, theme, or point of view clear and consistently maintained?
- **Support/Elaboration:** Are arguments and conclusions adequately supported and explained?
- **Organisation:** Is the logical flow of ideas clear and connected?
- **Conventions:** Are the standard English conventions (spelling, grammar, punctuation) properly followed? (Chapman, 1999)

However much one might agree with the apparent intent of the author, there is simply not enough definition in the criteria. What is 'clear', and to whom? How consistent is 'consistently'? How much support is adequate? Etc., etc., etc. Researchers have striven, therefore, to devise more precise, empirical techniques for evaluating Clear Writing — usually describing the outcome of that evaluation in the passive sense of *'Readability'*.

> *"There are over one hundred factors that affect how easy, or hard, a given document is to read and understand. These factors include sentence length, word choice, layout, tone, organisation, use of illustrations, and appeal to the reader. Readability formulas often only look at two or three of these factors — most commonly, the number of words in a sentence and the number of syllables in a word. ...Different tests for readability use different methods for choosing samples. It is important to follow the prescribed method for selecting a sample in order to get an accurate score."* (Osborne, 2000)

It is hardly surprising that with hundreds of Readability assessment formulae published there is considerable variation in the results. Our problem is confounded by the difficulty of accurately defining individual indicia; it is made even worse if we are attempting such definition for the purpose of computerisation. Even something as apparently simple as counting the number of words in a document can confound a multinational software developer. Checking this document at one stage of writing, Microsoft Word (Word Count) counted 15,218 words while the Grammar Checker in the same package counted 15,544. As I will demonstrate later, more complex evaluations generate greater differences with intricate and unpredictable variations.

These common measures of readability, it must be said, are not infallible and may create false impressions:

- English as a living language changes constantly. That which may be codified as *'good'* technical writing when the readability test is contrived may well be archaic when that test is implemented;
- Common readability tests such as the Flesch Readability Ease and the Gunning Fog Grade Level look for readability at the 50-75% comprehension level (Hargis, 2000, p.124) while the FORCAST Formula

Grade Level (which was devised for assessing U.S. Army technical manuals) was aimed at 35% (Johnson, 1998) . This is an amazing luxury to be afforded the writers of technical documentation whose readers have a right to expect to achieve near-100% comprehension of the subject, and fast! However, it does lead to the conclusion that technical documentation should, as Cranmer's *"Book of Common Prayer"* suggested, be in language *"understanded of the people"*.

> *"To facilitate higher comprehension, technical communicators need to use a lower readability target than the grade level that they expect typical users to have attained."* (Hargis, 2000, p.124)

- Unfortunately, many writers tend to overestimate the reading skills of the audience.

> *"Nearly half the adult population in the United States has low or limited literacy skills. ... there is not a one-to-one correspondence between the highest grade level completed and the literacy skills of an individual. While most adults in* [the US] *have completed high school, the average reading grade level is eighth to ninth grade. Like many other skills, reading skills atrophy if not used regularly, so people who read very little after they finish school will have correspondingly lower reading levels."* ((Merriman, Ades, and Seffrin, 2002, pp.130-1), see also (Mailloux, Johnson, Fisher and others, 1995, p.211))

- This gives further lie to the attitude referred to on page 7 that technical documentation should be esoteric. Readability may be reader-dependent — that which is comprehensible to the aficionado may be impenetrable to the tyro, yet both may have legitimate claim to the knowledge content of the document;

- Any grammar is composed of rules and the truism assures us that, *"Rules are for the obedience of fools and the guidance of wise men."* Authors frequently break grammatical rules to achieve impact, e.g. Hargis' article contains a subheading: *"Eschewing Gobbledygook"* (Hargis, 2000, p.127). That's sufficient of an eye-catcher to leave bruises on the eyeballs but plain language it is not! Although the phrase was intended to be an eye-catcher, and the plainness of the language is of little consequence, it would not enhance a readability mark in a standard test.

- Word selection, use of vocabulary to enhance precision, must be of considerable importance in establishing the ability of a technical document to convey information, yet it is almost impossible to factor into a readability formula. Consider Klare's example:

> *"The different meanings* [of the] *words 'wave' and 'waive' are, of course, obvious.* [Looking them up] *indicates that the meanings of 'wave' as a ripple on water, a curl in the hair or a hand signal are familiar to most readers with a fourth-grade education. However 'waive' as putting off or postponing, or as giving up a right, are familiar only to the majority of*

> *readers with a twelfth or thirteenth-grade education. Even the meanings of 'wave' as a line of attacking soldiers or as a motion of vibrating particles are… familiar only to the majority of readers with a twelfth-grade education."* (Klare, 2000, p.154)

Readability does, however, have an effect on speed of reading:

> *"We also found that the more readable versions produced an even stronger effect on reading speed and reader preference than on comprehension."* (Klare, 2000, p.158)

One might extrapolate that if the reader reads faster, for the same amount of time, then more will be read and, given the same percentage level of comprehension, more must be comprehended. One can, then, safely say that readability aids comprehension and thence that comprehension aids translation.

4.1 Computerised Evaluation of Written Assignments

Computer Assisted Analysis (CAA) of written work subdivides into assessment of content on the one hand and assessment of style on the other. It has been clearly shown that it is possible to effectively computerise both. Usually, either the student inputs answers via computer (online assessment) or enters them on a pre-printed test paper. Transferring the data to the computer is carried out via some form of scanner or other optical reader (Seale, [web page]). There is a third format in which digital copies of student prose assignments[9] are converted to ASCII text files and assessed by a computer program. Such a procedure appears to have considerable potential for both formative and summative assessment. Results from these assessments could be returned to the students. If students were also be provided with some instruction as to how these indicators can be used then the formative effect might be substantial. Individualised data could also be returned to the teaching faculty for the purpose of summative assessment. However, the major gain would be to provide a means to track the development of student technical writing skills (Section 2 discussed the importance of these skills to engineers).

It remains important to remember that any formulaic evaluation of written work remains open to criticism:

> *"While some people use readability tests to help them make their writing plainer, other people are fervently opposed to their use.*
>
> *"For Example, ten years ago the International Reading Association and the U.S. National Council of Teachers of English were advising members against uncritical use of readability tests to assess educational materials. At about the same time, two government reports in England validated the accuracy and reliability of the tests."* (Everything You Ever Wanted To Know About Readability Tests But Were Afraid to Ask., [web page])

Therein lies the key — the adjective *'uncritical'*. This study is undertaken to enable the tracking of <u>trends</u> in student technical writing. These tests should

not be used as an infallible guide for writing style:

> *"In all discussions of readability tests it should be remembered that they are designed 'for rating, not for writing'."* (Johnson, 1998, p.9)

Research constantly produces reminders that readability formulae are not designed as guides to writing style.

> *"A warning comes with all the formulas:* ***Do not write to the formula****. You cannot improve the readability of a passage simply by shortening words and sentences. You also have to attend to the tone, organisation, coherence, and design."* (Plain Language At Work Newsletter, 2004)

This does not mean that it is impossible to write on technical matters in a manner a novice could understand:

> *"Plain language is not 'dumbed down writing'. It is clear, conversational language designed to communicate effectively with a wide audience."* (Merriman et al., 2002, p.132)

> *"Psychologists may believe that their primary audience is other professionals. However, reports are distributed to parents, who are the primary consumers. ... it is entirely possible to write a psychological report at a comprehensible level and still communicate essential psychological information."* (Smith Harvey, [web page])

4.1.1 What the Formulae Cannot Do

Readability formulae are based on simple data: the number of words, the number of sentences, the number of monosyllabic words etc.

> *"Though crude, since* [Flesch Reading Ease] *is designed simply to reward short words and sentences, the index is useful. It gives a basic, objective idea of how hard prose is to wade through."* (Talburt, 1986, p.114)

Table 1: Stylistic Subtlety

The cat sat on your mat.	The cat on the mat sat.
The cat sat on the mat.	On the mat sat the cat.
On the mat the cat sat.	Sat: the cat on the mat.
Sat, on the mat: the cat.	Sat the cat on the mat?

Readability formulae do not take into account such subjective matters as font selection, layout or structure.

> *"... readability formulas do not take into account numerous variables that can influence the difficulty of a selection, such as legibility, motivation of the reader, learnability of the text, usability, relationship among words, sentences, and sentence parts, and the level of abstraction required of the reader by the material."* (Mailloux et al., 1995, p.225)

What may not be quite so obvious is that they do not take into account the writer's style. Consider the sentences in Table 1:

> *"Readability formulae will give the same value to each of these sentences even though the first example is probably the most readable. (It has a personal touch and a clear sequence of subject-verb-object which names the topic and then describes it.)*

> *"However readability formulae do distinguish very clearly between crisp and extended styles of writing. They will easily distinguish between: 'The cat sat on the mat' and 'The feline reclined on the axminster'."* (Johnson, 1998, p.9)

Neither do formulae allow for more subjective aspects of good writing, e.g. coherence:

> *"The most convincing ideas in the world. Expressed in the most beautiful sentences, will move no one unless those ideas are properly connected. Unless readers can move easily from one thought to another, they will surely find something else to read or turn on the television."* (Coherence: Transitions Between Ideas, [web page])

[9] At some later stage I will move into the field of student programming assignments.

In short, the output of readability formulae is a guide and only a guide:

"And spare us from the readability researchers who conclude that the reader needs 16·25 years of education to understand a document. Despite the statistical calculation, readability programs are not that accurate, and such precision has no basis in reality. 'Grade-16 reading level' is just another way of stating that the material is complex and average readers may find if very hard to read and understand."
{Hochhauser 1999 #1332 /ft ", p.22"}

4.1.2 Advantages of Automatic Analysis

Firstly and most obviously, CAA relieves the marker of the document from the drudgery of manual calculation of the statistics — counting words and syllables. This task, as indeed any marking work carried out by faculty, is generally regarded as one of the less exhilarating parts of the job:

"A daunting challenge for many teachers is the burden of evaluating student writing. Whether the assignment is a brief essay or a lengthy research paper, teachers are hard-pressed to read and grade student writing thoroughly and to provide the necessary discussion and feedback. In subjects other than English, the pressure to cover the subject matter discourages many teachers from giving extensive writing assignments … yet nearly all of them … say it is important or very important to assign such papers."
(Hurwitz & Hurwitz, 2004)

Other advantages present themselves, notably the consistency of purpose intrinsic to computer operation:

"Computers and artificial intelligence have been proposed as tools to facilitate the evaluation of student essays. In theory, computer scoring can be faster, reduce costs, increase accuracy and eliminate concerns about rater consistency and fatigue."
(Rudner and Gagne, 2001)

There are still further advantages:

"Administrative advantages:

- *Computerised marking is not prone to human error.*[10]
- *Saves time in terms of supervising and marking (including double marking) assessments.*
- *Reduction of printing costs, particularly when tests are updated or changed." (Seale, [web page])*

"[CAA systems are] *popular because they*

- *Reduce to simple formulae the complex work of writing*
- *Provide a convenient check and measure of the level of one's writing*
- *Possess the glitter of mathematical exactness*[11]
- *Can be calculated by word processing software." (Gunning Fog Index, 1999)*

"Some of the major advantages of using computers include reduction of work requirement, elimination of human error that is inherent in manual calculation, and lowering training requirements of the person assessing the readability." (Mailloux et al., 1995, p.222)

Finally, and obviously, computers (being mechanical) are objective and without any bias other than that introduced by the developer; do not suffer from lack of sleep; suffer interruptions gladly; never get bored; and are not annoyed by cold coffee. Their very non-humanity allows them to exhibit some scarce human virtues.

[10] As long as the program is correct ☹

[11] c.p. "The mathematical process for calculating RGL may give the impression of a greater degree of certainty than is warranted. Developing an ear for plain language is as important as testing printed materials for reading grade level." (Merriman et al., 2002, p.132)

4.1.3 Disadvantages of Automatic Analysis

There is, so it is said, no such thing as a free lunch and the advantages of CAA come at a price:

"Administrative disadvantages:

- *Implementing a CAA system can be costly and time-consuming.*
- *Staff who design and invigilate CAA need training in assessment principles and design, IT skills and examinations management.*
- *A high level of collaboration between all those involved in designing and implementing CAA is required."* (Seale, [web page])

Mechanical consistency, which provides many of the perceived benefits of CAA, also has its drawbacks:

"[CAA] *appears only to* [be] *applied to essays written in plain ASCII text. This artificial limitation in the application of the approach ignores the effect(s) on the marking of style enabled by text enhancements… Thus many style effects produced by modern word processors is* [sic] *wasted."* (Christie, [web page])

"The software developers acknowledge educators' concerns about their essay assessors' affects [sic] *on creativity and reflective thinking. But, they say, their software isn't meant to judge creative or sophisticated writing. Rather it is geared for expository essays on factual topics — papers describing how a psychologists theory works, for example, or relating the structure of the human heart…"* (Murray, 1998)

There is also the temptation to overemphasise the importance of the output of CAA:

"[CAA systems are limited because]

- *A low style of writing can result from slavish use of readability indexes — a monotonous succession of short sentences and simple words can make your writing dull and dull writing doesn't hold the reader's attention*
- *The indexes frequently give conflicting results."* (Gunning Fog Index, 1999)

Being formula-based, CAA cannot comment on some of the finer points of authorship:

"One feature of writing valued by writing experts that is not expressly represented in the current version of e-rater is coherence."[12] (Miltsakaki & Kukich, [web page])

Finally, any mechanical assessment system is vulnerable to manipulation when the submitting authors understand how it works:

"Critics argued that [CAA] *left the system vulnerable to cheating because students could artificially enhance their scores using tricks — they could simply write a longer essay, for example. …* [also, because CAA] *did not capture important qualities of writing such as content, organisation and style,* [it] *couldn't provide instructional feedback to students …"* (Kukich, 2000, p.22)

Despite these drawbacks, CAA has attracted growing support and acceptance.

4.2 Available Computerised Assessment Packages

CAA falls into two categories; assessment of style and assessment of content as discussed in section 4.1. In the United States, CAA is widely used by State Education and other authorities in situations where examinees number in (hundreds of) thousands — a practice which has grown steadily since the introduction of *Project Essay Grade* (PEG) in 1966. *Intelligent Essay Assessor* (IEA) followed

[12] "Coherence is product of many different factors, which combine to make every paragraph,

PEG in 1989 and Educational Testing Service's *Electronic Essay Rater* (e-rater) in 1999. Assessment sophistication increased with each development but while PEG concentrated on writing style the other two have moved to consider content as well (Rudner et al., 2001).

4.2.1 Assessment of Content

Scoring of the General Management Aptitude Test (on which much U.S. executive employment is based) is currently done by *e-rater* (Rudner et al., 2001) and IEA has been used to critique senior officers' papers for both the US Army and the USAF (Streeter, Psotka, Laham, & MacCuish, [web page]). Concerns as to the reliability of computerised assessment have been laid to rest:

> *"Two human readers agree 98% to 93% of the time. Using e-rater* [in early trials], *the rate of agreement between the computer and the human averaged 89%. ...A third reader is seldom needed. E-rater has been in production since February 1999 and has scored about 50,000 essays as of this writing. The agreement between e-rater and the human reader is between 88% and 95% (a percentage or two higher than between two human readers), with a mean average of 92%."* (Hedberg, 1999, p.5)

Assessment of the content of the essay poses particular problems, requiring the computer to be *'trained'*. Human markers are required to assess test documents and from their results the computer *'learns'* which essays should be related to which grading. In turn this requires a large body of *'training documents'*, which in turn means a large student cohort and a stable examination theme. These factors militate against essay assessment at this School using BETSY (Bayesian Essay Test Scoring sYstem)[13], a Windows-based package that requires a minimum 300 training documents for a particular topic at a particular level.

4.2.2 Assessment Of Style

Style analysis, on the other hand, presents fewer problems with many packages being readily available, including:

- *Readability Plus* which includes *Readability Calculations* (reporting Reading Grade Level (RGL) and raw data counts) and *Vocabulary Assessor* which can highlight words potentially difficult for the target audience (Readability Calculations, [web page]);
- *TextQuest* which has a test package which comes with no documentation, but which appears to attempt both content and style analysis[14];
- *Grammar Slammer* which requires the text to be pasted from the clipboard for the file under test — thus severely limiting file size — and which performs adjustable grammar and spelling checks[15]; and
- *Grammar Expert Plus* which performs spelling and grammar checks on text files and includes a simple text enhancer[16]. GXP also produces a Gunning Fog Index

every sentence, and every phrase contribute to the meaning of the whole piece." (Kies, 2004)

[13] See http://edres.org/betsy/
[14] See http://www.textquest.de
[15] See http://www.englishplus.com/
[16] See http://www.wintertree-software.com

calculation of readability (Wintertree Support Group, 2004).

- *TestStat 3.0*[17] allows browsing for a text file then produces various statistics including Flesch Reading Ease.
- *Microsoft Word Grammar Checker* is certainly the most regularly recommended package but is also regularly revealed to be somewhat unreliable. (This matter will be dealt with in a later chapter.).

When it comes to Software Engineering-specific document evaluation software ARM, the Automated Requirement Measurement Tool, was created by NASA in the mid-60s at the Goddard Space Flight Centre in Langley, VA (USA)[18]. Version 2·1 was produced in 1999. This Windows-based tool reads text files, producing a report for each file read, examining the linguistic voice of the document. For example, it seeks the presence and frequency of lexemes such as 'shall', 'must', 'is required to', 'will', 'should' etc. as evidence of the specificity of a requirements document. These lexemes may be individually deselected and reselected; other lexemes may be user-defined or deleted. This appears to be a useful training tool if very task-specific.

4.3 Reading Ease and Grade Level

Both Reading Ease and Reading Grade Level are calculated using simple weighted linear formulae producing a number. Flesch Reading Ease produces a number between 0 and 100; a higher score means easier comprehension. By contrast, Grade Levels are based on the US High School Grade System and a result of 7·0 from one or other formula would mean that a 7th grader (12-year-old) should be able to comprehend the text (Flesch Score, [web page]). These formulae and algorithms are readily available and sufficiently well-defined to be consistent to the third decimal place.

4.3.1 Flesch Reading Ease

$$206.835 - (1.015 \times ASL) - (84.6 \times ASW)$$

Formula 1: Flesch Reading Ease

where:

ASL = average sentence length (the number of words divided by the number of sentences)

ASW = average number of syllables per word (the number of syllables divided by the number of words)

[(Readability Formulas, [web page]), (Flesch-Kincaid Reading Measures, [web page]), (Nisus Writer Tips: Flesch Reading Ease, 2004), (Readability Test, 2004), (Caslon Analytics Profile: Online Readability, 2003)]

4.3.2 Flesch-Kincaid Grade Level

$$(0.39 \times ASL) + (11.8 \times ASW) - 15.59$$

Formula 2: Flesch-Kincaid Grade Level

where:

[(Readability Formulas, [web page]), (Flesch-Kincaid Reading Measures, [web page]), (Flesch-Kincaid Index,

[17] See http://www.lunerouge.com/freeware/freeware_e.htm

[18] See http://satc.gsfc.nasa.gov

2004), (Johnson, 1998, p.6), (Readability Test, 2004), (Reading in the Content Area, 2004), (Caslon Analytics Profile: Online Readability, 2003)]

4.3.3 Gunning Fog Grade Level

$$\left[\frac{W}{S}+\frac{T\times 100}{W}\right]\times 0.4$$

Formula 3 : Gunning Fog Grade Level

where:
W = number of words
S = number of sentences
T = number of difficult words (words containing three or more syllables) [(Readability Tests, [web page]), (Gunning Fog Index, 1999), (Johnson, 1998, p.5), (Readability Test, 2004), (Writing Tips, 1998), (Comm. 300 - Fog Index, (web page)), (Reading in the Content Area, 2004), (Caslon Analytics Profile: Online Readability, 2003)]

4.3.4 FORCAST Formula Grade Level

$$20-\frac{N}{10}$$

Formula 4 : FORCAST Formula Grade Level

where N = number of single-syllable words in a sample of 150 words [(Johnson, 1998, p.6), (Reading in the Content Area, 2004), (Gregory, [web page]), (Caslon Analytics Profile: Online Readability, 2003)]

4.4 Density of Language

Density is expressed as a ratio of the number of unique words (or specified words) to the total number of words in a document. *Analyse* assesses the document for two densities: (1) vocabularic density, which looks at all the words in the document and (2) lexical density, which calculates the frequency of appearance of lexemes specified by a standard glossary or lexicon.

4.4.1 Vocabularic Density

VD provides an insight into the author's command of the language. Introducing a new word to the document requires (1) that the author knows the word (has a broad vocabulary) and (2) that the author has it on recall (has an easy familiarity with the vocabulary).

Vocabularic density is a simple ratio but it is mentioned in a two forms:

$$W/U$$

Formula 5 : VD

where:
U = number of unique words in the document and
W = total number of words in the document. (Lexical Density Test, [web page])

and

$$U/W$$

Formula 6 : Inverse VD

(Simpson, 2002)

VD is calculated by creating a lexicon of all the words in the document and counting multiple appearances. It is then possible to establish the ratio between the number of words that only appear once and the total word count

for the document. Obviously for a small document the ratio

$$TotalWords/UniqueWords$$

may approach 1 — e.g. in a document of ten words it would be easy to have each word appearing once. On the other hand, in a large document (Total Words→∞) the ratio will approach infinity because the author's vocabulary is finite. Figure 4 shows the relationship between VD and document size for a range of documents including magazine and journal articles, textbooks, literary works and technical manuals. However, the asymptotic nature of Vocabularic density is more clearly demonstrated by graphing the inverse as in Figure 5.

This means that, to obtain any useful information from Vocabularic Density (VD) it is necessary to analyse a document of at least 2,000 words while for anything above 50,000 words variation of VD tends towards the negligible. Fortunately, these parameters cover the expected size of major student assignments.

Note that VD is difficult to sustain at a density >0·075 in a document over 5,000 words. For example, the novel *Moby Dick* has a new word every 12 (VD=0·083) in a total of some 250,000 words which is considered exceptional. (Simpson, 2002)

4.4.1.1 OVIX

Recently developed by Professor Tor Hultman at Lund University in Sweden, OVIX normalises the VD ratio using

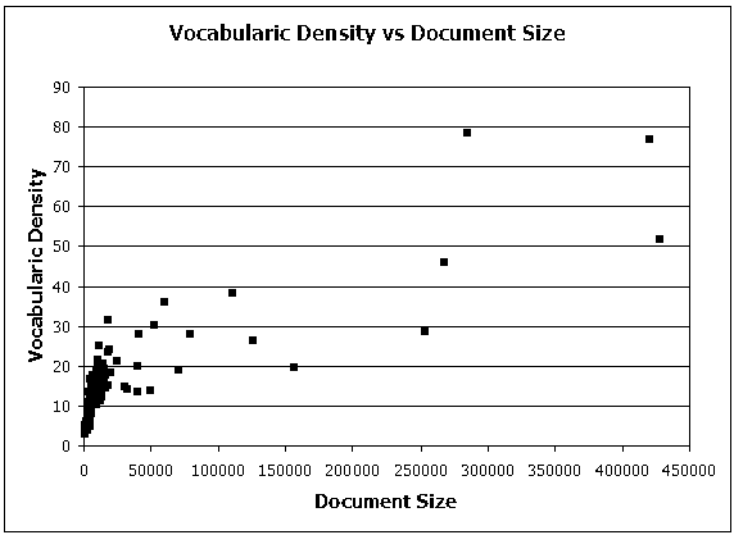

Figure 4: Vocabularic Density vs Document Size

natural logarithms to flatten the abovementioned asymptotic nature of the W:U curve. Several formulae were devised by Prof. Hultman; Formula 7 is the formula applicable to computerised calculation of VD.

$$-1\Big/\left[\frac{\ln(\ln(v)/\ln(n))}{\ln(n)}\right]$$

Formula 7: OVIX

where:

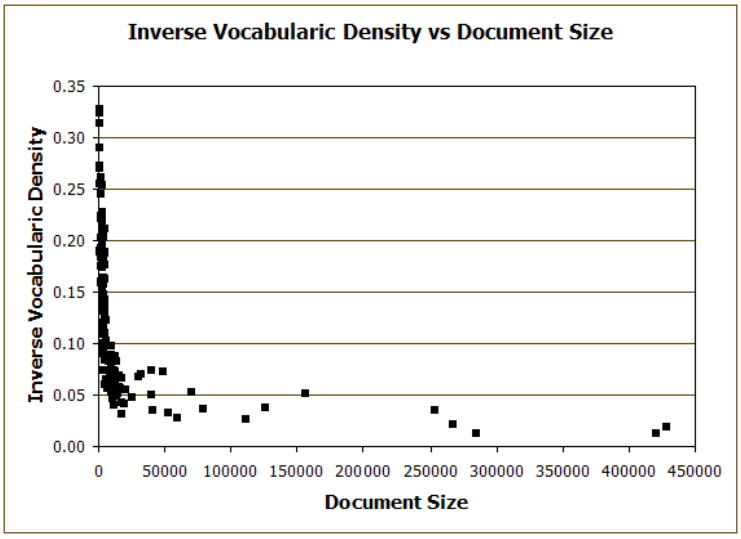

Figure 5: Inverse Vocabularic Density vs Document Size

ln = natural logarithm,
v = the number of unique words and
n = the total number of words.
(Hultman, 1993, p.62)

4.4.2 Lexical Density

Analyse uses a standard lexicon — default is (IEEE Standard Glossary of Software Engineering Terminology, 1999) — for comparison with the text. Quite simply, *Analyse* tests each sentence for the presence of lexemes from the lexicon and counts their appearances. This count is used in a calculation similar to that for VD:

$$WordCount / Lexemes$$

to provide an indicator of similar use to VD but focussed on the author's familiarity with the standard terminology of the discipline.

4.5 Requirement Measurement Functionality

NASA's Goddard Space Flight Center developed the Automated Requirement Measurement (ARM) tool to monitor the software requirements documents produced throughout NASA[19]. ARM tests sentences for occurrences of certain indicators and publishes a report for the document listing the occurrences in their context. Six types of indicator are sought: Imperative, Continuance, Directive, Option, Weak Phrases, and Incompletes. Default examples of the imperative type are: 'shall', 'must', 'is required to', 'are applicable', 'are to', 'responsible for', 'will' and 'should'. (Software Assurance Technology Center, 1996, p.8) Emphasising the presence of these indicators in a draft version of the document provides the author with the opportunity to use alternative expressions. This, in turn, provides the opportunity for *'tightening up'* the wording of the document to minimise *'ambiguity, inaccuracy and inconsistency'*. Since the terminology in Design, Maintenance or User documentation should be as precise as that in a Requirements document, this functionality should provide useful information for document revision in all phases of the Software Development Lifecycle.

4.6 Using Available Assessment Software

One consistent problem with using any of these Windows-based packages is that they deal with one file at a time. Students checking their own work would find these products useful but they are no good at all for a researcher looking for trends across the whole student cohort. In this case, the need is for a product that will successively analyse a collection of files and generate data in a format suitable for input to analysis packages such as Excel, SPSS, Access or MySQL. None has been found to date; that left the option of creating one — I have and I called it *Analyse*.

[19] See http://satc.gsfc.nasa.gov

5 A Final Word on Formulae

This from George R. Klare, *eminence grise* of the readability community:

Something has already been said of what formulas cannot be expected to do — their limitations. Formulas have been criticised over the years, for example, because of the mistaken assumption that they were designed to measure all the important aspects of writing. On the other hand, some users have accepted formula scores uncritically even where there is evidence against their acceptance. So at least brief notice of the limitations of formulas seems called for.

"First, *formulas measure only one aspect of writing — style. Perhaps certain formulas, especially those concerned with abstractness of words or analysis of ideas, approach a measure of content, but they only touch on this indirectly. Yet work in readability constantly shows that unless content is interesting and relevant to the reader's need, it will not make much difference whether it is readable or not, since it will seldom be read. Nor is content the only concern. Formulas do not touch on organisation, word order, format, or imagery in writing; they do not take into account the differing purposes, maturity, and intelligence of the readers.*

"Second, *formulas measure only one aspect of style — difficulty. Other aspects of style are important, as any literary critic can point out. Those that come to mind are the dramatic effectiveness of the writing, or its ability to create a mood; writing may sway opinions or satirise public life, but a formula will not measure its effectiveness in doing so.*

"Third, *formulas do not even measure difficulty* perfectly. *Formulas appear to give scores accurate to, or even within, one grade level. Yet actually they are seldom this accurate. The factors just mentioned frequently play a part in determining whether a piece of writing is at a particular level of difficulty in any given circumstance. And beyond these, a formula score may be inaccurate due to errors in sampling or in application.*

"Fourth, *formulas are not measures of good style. Clearly, a poor piece of writing may sometimes be poor because it has a low readability score, but it would not necessarily be good if it had a good readability score. Therefore, the temptation simply to substitute an easy word for a hard one, to cut sentences in half, or to eliminate prepositional phrases even when they are needed may produce exactly the opposite of the desired result.*

"Thus style difficulty is shown to be only one characteristic of a piece of writing. It is often an important one, however, as shown by its relationship to reading speed, acceptability, understanding, and learning. If formulas are thought of as efficient predictors of difficulty, more accurate in prediction than individual writers most of the time, that is all that should be expected. At present, at least, content and all the other elements of writing, plus aspects of style other than difficulty, are not taken into account in formula scores. Formulas can be highly useful to the professional communicator if these limitations are kept in mind. (Klare, 1963, pp.120-121)

Perhaps the only accurate way to assess the level readability of a document is to have it read by people at a range of reading levels.

Reference List

"*Writing Tips*", University of Minnesota (USA) (Published: 1998) URL: http://www.fpd.finop.umn.edu/groups/ppd/documents/information/writing_tips.cfm, Accessed: June 8, 2004.

IEEE Standard Glossary of Software Engineering Terminology (1999): Std. No. IEEE Std 610.12-1990. New York, NY (USA): IEEE Computer Society

"*Gunning Fog Index"* (Published: 1999) URL: http://www.soar.gcps.k12fl.us/pdfs/gunning.pdf, Accessed: June 29, 2004.

"*Caslon Analytics profile: online readability"*, Caslon Analytics (Published: 2003) URL: http://www.caslon.com.au/readabilityprofile1.htm, Accessed: July 14, 2004.

"*Nisus Writer Tips: Flesch Reading Ease"*, Nisus Software Inc. (Published: 2004) URL: http://www.nisus.com/NisusWriter/Suppor/Tips, Accessed: June 8, 2004.

"*Readability Test"*, Juicy Studio (Published: 2004) URL: http://www.juicystudio.com/fog, Accessed: June 8, 2004.

"*Reading in the Content Area"*, Washington County Public Schools, MD (USA) (Published: 2004) URL: http://www.alt.wcboe.k12.md.us/mainfold/technolog/instruct/msde07/module14/mod14_activity_c.htm, Accessed: July 14, 2004.

"*Plain Language At Work Newsletter"*, Impact Information Plain-Language Services (Published: 2004) URL: http://www.impact-information.com/impactinfo/newsletter/plwork08.htm, Accessed: July 21, 2004.

"*Flesch-Kincaid Index"* (Published: 2004) URL: http://www.usingenglish.com/glossary/flesch-kincaid-index.html, Accessed: June 8, 2004.

"*Coherence: Transitions between ideas"* (Published: [web page]) URL: http://webster.commnet.edu/grammar/transitions.htm, Accessed: March 5, 2004.

"*Comm. 300 - Fog Index"*, University of Arizona (USA) (Published: (web page)) URL: http://www.ic.arizona.edu/~comm/mary/fogindex.html, Accessed: June 8, 2004.

"*Everything You Ever Wanted To Know About Readability Tests But Were Afraid to Ask."* (Published: [web page]) URL: http://www.gopdg.com/plainlanguage/readability.html, Accessed: February 5, 2004.

"*Flesch-Kincaid Reading Measures"* (Published: [web page]) URL: http://www.msresource.com/misc/Flesch-Kincaid.htm, Accessed: June 8, 2004.

"*Flesch Score"* (Published: [web page]) URL: http://www.measurementexperts.ord, Accessed: June 8, 2004.

"*How to Write in Plain English"*, Plain English Campaign (Published: [web page]) URL: http://www.plainenglish.co.uk/plainenglishguide.html, Accessed: February 5, 2004.

"*Lexical Density Test"* (Published: [web page]) URL: http://www.usingenglish.com/glossary/lexical-density-test.html, Accessed: April 27, 2004.

"*Readability Calculations"* , Micro Power and Light Co. (Published: [web page]) URL: http://www.micropowerandlight.com/rdplus.html, Accessed: February 9, 2004.

"*Readability Formulas"*, Memphis University (USA) (Published: [web page]) URL: http://csep.psyc.memphis.edu/cohmetrix/readabilityresearch.htm, Accessed: April 27, 2004.

"*Readability: Frequently Asked Questions"*, Department of Administrative Services, OR (USA) (Published: [web page]) URL: http://www.oregon.gov/DAS/readability.shtml, Accessed: June 28, 2004.

"*Readability Tests*" (Published: [web page]) URL: http://developer.gnome.org/documents/style-guide/x3568.html, Accessed: June 8, 2004.

Australian National Training Authority, "*At A Glance: Defining Generic Skills*", NCVER (Published: 2003) URL: http://www.ncver.edu.au/research/proj/nr21026.pdf, Accessed: October 8, 2003.

Australian Technology Network, "*Developing Written and Oral Communication Skills in Engineering Students*" (Published: [web page]) URL: http://cea.curtin.edu.au/ATN/rmit_eng_report.html, Accessed: October 3, 2003.

Barthel, A., University of Technology, Sydney, NSW; "*Academically Speaking: New identities or old realities?*" (Published: [web page]) URL: http://learning.uow.edu.au/las2001/unrefereе3d/barthel.pdf, Accessed: July 14, 2004.

Brown, J.S., Collins, A, & Duguid, P, "*Situated Cognition and the Culture of Learning*" (Published: 1989) URL: http://www.ilt.columbia.edu/ilt/papers/JohnBrown.html, Accessed: May 30, 2003.

Chapman, C., Illinois State Department of Education; "*Authentic Writing Assessment*", Practical Assessment, Research and Evaluation (Published: 1999) URL: http://pareonline.net, Accessed: February 8, 2004.

Christie, J.R., The Robert Gordon University, Aberdeen (SCO); "*Automated Essay Marking - for both Style and Content*" (Published: [web page]) URL: http://www.lboro.ac.uk/service/fli/flicaa/conf99/pdf/christie.pdf, Accessed: February 5, 2004.

Deacon, A.B., University of Wisconsin-Milwaukee; "*A Return to Computerized Essay Scoring: What's at Stake?*", UWM (Published: [web page]) URL: http://www.uwm.edu/~abd/scoring.html, Accessed: 19, 2004.

Flesch, R., "*How to Write Plain English*", Canterbury University (NZ) (Published: [web page]) URL: http://www.mang.canterbury.ac.nz/courseinfo/AcademicWriting/Flesch.htm, Accessed: April 27, 2004.

Forward, A.; *"Software Documentation - Building and Maintaining Artefacts of Communication"*. (Published: 2002) University of Ottawa - Ottawa (CA).

Friere, P. (1972). Pedagogy of the Oppressed. Harmondsworth (UK): Penguin Education.

Gersting, J. L., & Young, F. H.; *"Shall We Write?"*. SIGCSE Bulletin , (June, 2001) 33(2):18-19.

Gregory, E., Texas Cooperative Extension; "*Readability*", Agricultural Communications (Published: [web page]) URL: http://agcommwww.tamu.edu/market/training/power/readabil.html, Accessed: July 14, 2004.

Hallgren, C. (1986) *"Factors Affecting Readability"*. Proceedings of the 4th Annual International Conference on Systems Documentation. (pp. 108-109). New York, NY (USA): ACM.

Hargis, G.; Senior Software Engineer, (IBM Corporation): *"Readability and Computer Documentation"*. ACM Journal of Computer Documentation, (August, 2000) 24(3):122-131.

Hedberg, S. R.; *"Computers scoring GMAT essays? Impossible! Or is it?"*. IEEE Intelligent Systems, (May, 1999):5-7.

Hofstede, G.; *"Cultural Dimensions in Management and Planning"*. Asia-Pacific Journal of Management, (1984) 1(2):81-99.

Huckin, T. N., & Olsen, L. A. (1991). *"Technical Writing and Professional Communication for nonnative speakers of English"*. New York, NY (USA): McGraw-Hill Inc.

Hultman, T. (1993). Hur gick det med OVIX? In U. Teleman (ed.), Språkbruk, grammatik och språkförändring. (pp. 55-64). Lund (SWE): Lund University.

Hurwitz, N., & Hurwitz, S, "*Words on Paper"*, American School Board Journal (Published: 2004) URL: http://www.asbj.com/2004/03/0304coverstory.html, Accessed: June 29, 2004.

IEEE/ACM Joint Task Force on Computing Curricula (2003); New York, NY (USA): IEEE-CS/ACM.

IEEE/ACM Joint Task Force on Computing Curricula: *"Computing Curriculum - Software Engineering"* (2004): Report No. Public Draft 3.1; New York, NY (USA): IEEE-CS/ACM.

IEEE-CS/ACM Joint Curriculum Task Force, "*Computing Curricula 2001 (Final Report)"* (Published: 2001) URL: http://computer.org/education/cc2001/, Accessed: December 23, 2001.

Johnson, K., "*Readability"* (Published: 1998) URL: http://www.timetabler.com/reading.html, Accessed: June 8, 2004.

Kies, D., College of DuPage; "*Coherence in Writing"*, Hypertext Books (Published: 2004) URL: http://papyr.com/hypertextbooks/engl_101/coherent.htm, Accessed: March 3, 2004.

Klare, G. R. (1988). The Formative Years. In Zakaluk & Samuels (Eds), Readability: Its past, present, and future (pp. 14-34). Newark, DE (USA): International Reading Association.

Klare, G. R. (1963). Useful Information for Communicators. Chapter One In The Measurement of Readability (pp. 11-25). Ames, IA (USA): Iowa State University Press.

Klare, G. R.; (Ohio University): *"Readable Computer Documentation"*. ACM Journal of Computer Documentation, (August, 2000) 24(3):148-168.

Kukich, K.; (Educational Testing Service): *"The Debate on Automated Essay Grading"*. IEEE Intelligent Systems, (September, 2000):22-27.

Levine, L., Pesante, L. H., & Dunkle, S. B.: *"Technical Writing for Software Engineers"* (1991): Report No. SEI-CM-23; Pittsburg, PA (USA): CMU/SEI., Accessed: September, 2003.

Mailloux, S. L., Johnson, M. E., Fisher, D. G., & Pettibone, T. J.; (Kenai Care Center, AK (USA)): *"How Reliable is Computerized Assessment of Readability?"*. Computers in Nursing, (September, 1995) 13(5):221-225.

Mayer, E.; Chair, Committee to advise the AEC and MOVEET: *"Putting General Education to Work: The Key Competencies Report"* (1992); Melbourne (VIC): The Australian Education Council and Ministers for Vocational Education, Employment and Training.

Merriman, B., Ades, T., & Seffrin, J. R.; *"Health Literacy in the Information Age: Communicating cancer information to patients and families"*. CA - A Cancer Journal for Clinicians, (May, 2002) 52(3):130-133.

Miltsakaki, E., & Kukich, K, "*Automated Evaluation of Coherence in Student Essays"*, University of Pennsylvania (Published: [web page]) URL: http://www.ling.upenn.edu/~elenimi/lrecpaper.pdf, Accessed: February 12, 2004.

Mogyrodi, G. E.; (Bloodworth Integrated Technology Inc.): *"What is Requirements-Based Testing?"*. Journal of Defense Engineering, (March, 2003):12-15.

Murray, B. (August, 1998). *"The latest techno tool: essay-grading computers"*. APA Monitor, 29(8). URL: http://www.apa.org/monitor/aug98/grade.html Accessed: July 21, 2004

Osborne, H. (Health Literacy Consulting). (December, 2000). *"In Other Words... Assessing Readability... Rules for Playing the Numbers Game"*. On Call. URL: http://www.healthliteracy.com/oncalldec2000.html Accessed: September 9, 2003

RMIT, Royal Melbourne Institute of Technology; "*Written and Oral Communication"* (Published: [web page]) URL: http://www.clt.uts.edu.au/RMIT.engineering.html, Accessed: October 3, 2003.

Rudner, L., & Gagne, P.; (University of Maryland, College Park, MD (USA)): *"An Overview of Three Approaches to Scoring Written Essays by Computer"*. Practical Assessment, Research and Evaluation, (December 12, 2001) 7(26).

Seale, J., "*Computer Assisted Assessment: Using CAA to support student learning"* (Published: [web page]) URL: http://www.swap.ac.uk/elearning/develop6.asp, Accessed: February 5, 2004.

Simpson, Z.B., "*Vocabulary Analysis of Project Guttenberg"* (Published: 2002) URL: http://www.mine-control.com/zack/guttenberg/, Accessed: March 3, 2004.

Smith Harvey, V., University of Massachusetts, Boston, MA (USA); "*Improving Readability of Psychological Reports"* (Published: [web page]) URL: http://www.psychpage.com/learning/intell/harvey.html, Accessed: July 21, 2004.

Snoke, R., & Underwood, A. (1999) *"Generic Attributes of IS Graduates: An Australian IS Academic Study"*. Proceedings of the 10th Australasian Conference on Information Systems.

Automated Requirement Measurement Tool (Version 2.1), Software Assurance Technology Center; SATC (1996).

Streeter, L., Psotka, J, Laham, D, & MacCuish, D, "*The Credible Grading Machine: Automated Essay Scoring in the DoD"*, Knowledge Technologies (Published: [web page]) URL: http://www.knowledge-technologies.com/papers/essayscoring.pdf, Accessed: February 5, 2004.

Talburt, J., University of Arkansas at Little Rock . (1986) *"The Flesch Index: An Easily Programmable Readability Analysis Algorithm"*. Proceedings of the 4th Annual International Conference on Systems Documentation. (pp. 114-122). New York, NY (USA): ACM.

Teichler, U.; *"Higher education policy and the world of work: Changing conditions and challenges"*. Higher Education Policy, (1999) 12:285-312.

Wheeler, T.M., "*Writing for Readability"* (Published: [web page]) URL: http://www.seeseeeye.com/pages/learn/whitepages/readability/WritingforReadability.pdf, Accessed: July 21, 2004.

Wilson, B., Jonassen, D., & Cole, P. (1993). Cognitive Approaches to Instructional Design. In G. M. Piskurich (Ed.), The ASTD handbook of instructional technology (p. 21.1-21.22). New York, NY (USA): McGraw-Hill.

Wintertree Support Group (support@wintertree-software.com). (2004, March 8). Readability Test in GXP. E-mail to Rick Duley (rick@eng.murdoch.edu.au).