# Wycliffe Associates EasyEnglish: Challenges in Cross-Cultural Communication

**Robert G. Betts** 

Wycliffe Associates (UK) 7 Conwy Street Rhyl Denbighshire LL18 3ET United Kingdom rbetts@btinternet.com

#### Abstract

A formally defined subset of English known as EasyEnglish is described. Primarily designed for a culturally diverse audience of second-language English speakers, EasyEnglish seeks to clarify meaning through restrictions on vocabulary and grammar, whilst preserving intellectual content. EasyEnglish has been developed by Wycliffe Associates (UK), an organisation that produces Biblical materials for a worldwide audience. Special features of EasyEnglish include: (1) two levels of lexical restriction (approximately 1,200 words and 2,800 words respectively); (2) constraints on the range of meanings allowed for permitted terms; (3) logical reformulation and application of grammatical structures designed to maximise comprehensibility; and (4) application of established techniques for translation across cultural boundaries. Printed and computerised tools for writing are briefly outlined, together with future developments. The efforts we make in making the EasyEnglish texts clear to people from a wide diversity of cultures and who speak a wide range of mother tongues may contribute significantly to the wide acceptance of our material; valid criticisms of EasyEnglish are also pointed out. Three particular current challenges are: selection of corpora from which suitable vocabulary can be drawn; fine-tuning the grammatical rules; and obtaining feedback from users.

#### 1 Introduction

Wycliffe Associates (UK) produces Biblical materials for pastors, teachers of English, Bible translators, etc., worldwide. Our audience may have limited English as a second or subsequent language, or they may be training or communicating to such people.

Our main thrust currently is production of Bible translations and commentaries at two levels of simple English. We originally conceived the simpler level as a Bible translation aid for mother-tongue translators who understand some English. But it has wider use for those with limited English or reading difficulties. The more advanced level is designed for leaders and teachers who may have had little or no theological training. EasyEnglish translations may also serve as texts for those learning English.

We currently distribute our material by CD and by Website download (with approximately 250,000 hits and 100,000 downloads per year from 140 countries). Only a small proportion of our audience owns a computer able to access the Web or read CDs. But increasingly there are Internet cafés even in quite small towns in many parts of the developing world. This enables rural people to access and download files. These people who use computers are 'gateways' to a much larger audience. For example, a church leader may use the material to teach other leaders or the whole congregation; schoolteachers may use the material with their classes.

We are diversifying into production of other materials (including those for the learning disabled), and this diversification will increase as the Bible translation process draws to its completion.

## 2 What is EasyEnglish?

EasyEnglish is a formally defined subset of standard English, with restricted vocabulary and grammar. It is different from IBM's EasyEnglish system (now called EasyEnglishAnalyser), described by Bernth (1997).

We see our work of producing materials in EasyEnglish as a translation process. A special challenge is to express complex or abstract ideas in simpler words and grammatical structures without losing meaning. There are four basic elements:

- 1. Lexical restriction
- 2. Simplified grammatical structure
- 3. Logical restructuring to optimise comprehensibility
- 4. Use of well-established Bible translation techniques for communicating across cultural boundaries.

## 2.1 Restrictions in Vocabulary

There are two levels of EasyEnglish:

- 1. Level A uses approximately 1,200 words (3,000 inflections). It assumes a working knowledge of English as a second language sufficient to cope with most social and work situations.
- 2. Level B uses approximately 2,800 words (7,000 inflections). It is directed at those with an intermediate level of proficiency in English.

The vocabulary is restricted in two ways: the choice of words, and the choice of meanings for each word.

#### **Choice of Words**

EasyEnglish currently has a restricted vocabulary based on the 'Cambridge English Lexicon' (Hindmarsh, 1980). This was a lemmatized compilation of high-frequency words in English. We are now developing a new vocabulary with the help of a frequency list based on a larger published English corpus, with our own definitions of these words. The list is being adjusted by experienced Easy-English writers to allow for the more everyday words we need when writing for (often) Third-World readers, and to reject some high-frequency (often technological) words that we will probably never need. Other cross-cultural factors influence our choice of terms. In particular, we use few abstract nouns (for example: 'faith', 'strength', 'anger', 'peace', 'ability'), since some mother languages spoken by potential users of our material have relatively few nouns of this type.

A major challenge is developing a corpus for Although our present output is EasyEnglish. mainly Bible translations and commentaries, we are diversifying and cannot restrict ourselves to Biblical corpora. There are a number of word frequency counts available, for example: the Thorndike-Lorge semantic count (1938), the Brown Corpus (Francis and Kucera, 1982), the LOB Corpus (Johansson and Hofland, 1989). More modern corpora include the CoBuild corpus (http://titania.cobuild.collins.co.uk) and the British National Corpus (BNC); a lemmatised frequency list of the BNC has been prepared by Adam Kilgarriff (http://www.hit.uib.no/corpora/ 1996-2/0066.html). But these corpora are not suitable for us without considerable adaptation. They reflect their source material, which may differ considerably from the vocabulary used by our target audience. Two problems stand out:

Firstly, our audience is to a large extent from the Third World, often rural and inhabiting a world very different from a western technological environment. The set of objects and ideas that they use daily are different from the set with which we are familiar, and on which these frequency corpora are largely based.

Secondly, our audience is also very diverse. We distribute our materials mainly via the Internet, and the approximately 100,000 downloads in 2002 went into over 140 countries. We aim to supply materials to those in cities as well as remote rural locations - from South America, through Europe to the Far East. No single corpus of vocabulary would serve everyone equally well. We have to have a 'happy medium' - but where to centre that happy medium is difficult to decide. Do we bias it toward rural Africa (for example), the slums of Rio or the business district of Bangkok?

Corpora based on published material taken from print or the Internet are, in our view, inadequate for our purposes. They are unlikely to reflect the English vocabulary used by the bulk of our audience. We need corpora derived from collections of spoken as well as written everyday English from different parts of the world. One approach might be to obtain literature in simple English (such as texts used in primary/secondary education) in many of our key user locations - Africa, India, South-East Asia, South America.

## **Total Vocabulary Size**

Having obtained a corpus, what words do we select? In general, the most frequent - but how far down the frequency rankings should we go? In other words, how big should our selected corpus be? There is no hard and fast rule. Experience has shown that we can successfully translate the Bible with a vocabulary of 1,200 words without significant loss of meaning; use of the 2,800 words at Level B can provide a translation with some degree of style and sophistication. West (1950) states (in relation to story-telling) that at 750 words one can tell fairy stories well and an adventure story with difficulty; at 1,100 words one can tell adventure stories well, but in rather a bald manner; at 1,700 words one can tell any plot, preserving much of the original style; and that a vocabulary of 2,000 words is sufficient for anything, and more than sufficient for most things.

#### **Numbers of Meanings**

Selection of words is only part of the process of building a working corpus. Four aspects of how EasyEnglish deals with meanings may be mentioned.

Firstly, English words often have many meanings. For example, 'fair' can mean 'beautiful', 'blond', 'unbiased', 'reasonably good', 'favourable', 'market', 'amusement show' and 'commercial exhibition'. Some of these meanings may be misunderstood by or opaque to non-native English speakers. We must decide what meanings of words are allowed. To do this well requires considerable understanding of the target audience's use of English. EasyEnglish prefers primary meanings (i.e. the sense that first comes to mind when the term occurs in isolation) - for 'fair', Level A EasyEnglish allows only 'unbiased'. Even a phrase like 'coming to see you' is an extended use of the verb 'to see'; it actually means 'coming to meet you'. This use of 'to see' can confuse people with limited English as their second language. Accordingly, EasyEnglish avoids it at Level A.

Secondly, and consistent with its effort to favour primary meanings, EasyEnglish avoids idioms, metaphors and figures of speech (such as hyperbole, rhetorical questions, litotes, and euphemisms).

Thirdly, many words are used as different parts of speech. Non-native English speakers can find this confusing. For example, 'wrong' can be used as a noun, verb or adjective. EasyEnglish allows 'wrong' only as an adjective.

A final point is that the '-*ing*' inflection of the verb may be either a participle (an adjectival or a verbal form) or a gerund (functioning as a noun). The '-*ing*' form has many features unique to English and can confuse second-language English

speakers. EasyEnglish tries to use this grammatical form only for the present or past continuous tense (*'the wind is roaring', 'the wind was roaring'*). Where possible, it avoids its use as an adjective (*'the roaring wind'*) and a noun (*'the roaring of the wind'*).

#### 2.2 Restricted Grammar

Lexical considerations, however, are not the biggest concerns when designing an English system for second-language speakers in diverse cultural settings. The following rather amusing example demonstrates this. It was quoted (Anon., 2002) as an example of inappropriate language for nonnative English speakers. The Royal London Hospital in Whitechapel has been for centuries in a neighbourhood where many new arrivals in the UK first settle. For these people, English may well be a third or fourth language. The outpatients department displayed this notice for people waiting for a blood test:

If you are attending another clinic and having your blood taken with your yellow book when there is not an anticoagulation clinic going on downstairs please check with the phlebotomist and take a ticket as you would normally.

The problem here is not only (or even primarily) about vocabulary. Probably only two words (*'anti-coagulation'* and *'phlebotomist'*) should be replaced or explained. The real problem is grammar. The sentence is too long, the train of thought convoluted, and it seems to assume prior knowledge of certain facts (i.e. there appears to be implicit information). So restricting the vocabulary is not enough. Attention must be paid to simplifying grammar and sentence structure.

EasyEnglish's grammatical structure is designed with one goal in mind: clarity. The structure is based on work done by Wycliffe Associates member Karen Bennett (Bennett, undated). Bennett developed grammatical rules for Easy-English largely from her own experience as someone whose mother tongue was not English, and also as a teacher of English. She found no single source of advice to call upon, although (for example) the Plain English Campaign's materials (www.plainenglish.co.uk) offered significant help.

Bennett studied sample English texts to see what made them complex. For example, she found that Readers' Digest texts (designed as easyreading texts) often used complex sentence structures; for example, sentences that contain multiple prepositional clauses. She found that complexity is determined more by the number of 'idea units' or 'units of meaning' per sentence than by vocabulary. So she developed a simplified grammatical system based on these findings and on her own experience as both a second-language speaker and a teacher of English. Some of the EasyEnglish grammatical restrictions are:

- 1. **Length** We limit sentence length to 20 words maximum, and aim for 8-12 words. Paragraph length is limited to 150 words.
- 2. Sentence structure We allow a maximum of two finite clauses per sentence (this is a difficult area and is being reviewed). No more than two prepositional phrases are permitted (e.g. 'they ran to the shop in the village' is allowed; 'they ran to the shop in the village before sunset' is not). We generally limit embedding to one phrase; thus the sentence 'Along the coast there are picturesque little harbours and old and ruined castles perched on precipitous rocky cliffs' would be simplified in EasyEnglish by division into two or more sentences.
- 3. **Passives** EasyEnglish avoids passives (with a very few exceptions considered both essential, and clear to our audience, such as '*called*' and '*born*'). These are relatively complex forms. In addition, there is an immense variety in the use of the passive voice in languages. Some languages (e.g. including many of the around 850 languages spoken in Papua New Guinea) do not use passives. Mother-tongue speakers of these languages may well find the English passive difficult to grasp.
- 4. Ambiguous grammatical forms. For example, EasyEnglish avoids pronouns that can refer back to more than one noun. The use of the genitive is restricted to forms that make relationship between the the terms unambiguous. For example, 'the city of Thessalonica' might be interpreted as a city in a district called Thessalonica. A better alternative is 'the city called Thessalonica'.

#### 2.3 Logical simplicity

Even texts that use the EasyEnglish vocabulary and obey its grammatical rules can still display ambiguity. The sentence '*He hit the man with the umbrella*' displays functional ambiguity. It conforms to EasyEnglish vocabulary and grammar, but still needs rewriting.

On a broader scale, EasyEnglish asks for a logical flow of ideas easily grasped at the first

reading. The original text being translated into EasyEnglish may require a slight re-ordering to achieve this. EasyEnglish writers are encouraged to identify the basic idea units in a complex sentence or paragraph and arrange them in logical order before attempting translation. This helps to identify problem areas and eases translation. An example is:

Even the glorious loneliness of the Highland's wonderful landscape of loch, moor and mountain is largely a product of the 'Clearances' of the 18th and 19th centuries, which caused so much hardship and suffering.

There are no less than six idea units here:

- 1. The landscape of the Highlands consists of loch, moor and mountain.
- 2. This is a wonderful landscape.
- 3. The landscape is gloriously lonely.
- 4. The loneliness is largely a product of the Clearances.
- 5. The Clearances occurred in the 18th and 19th centuries.
- 6. The Clearances caused much hardship and suffering.

Identifying and arranging the elemental idea units in this way enables the EasyEnglish translator to reassemble them in a series of short, simple sentences conveying a logical flow of ideas that builds the readers' knowledge step by step. This process is as essential as adhering to the rules of grammar and vocabulary. The above example might result in this EasyEnglish translation (*'Clearances'* and *'moors'* are outside the EasyEnglish vocabulary, and are thus explained):

The Highlands of Scotland consist of lakes, mountains and moors. The moors are flat empty lands where no trees grow. This land is wonderful and magnificent because it is so empty. However, many people once lived there. But in the 18th and 19th centuries the owners of the land forced these people to leave. These people suffered many difficulties and troubles. People call these terrible events 'the Clearances'.

Note that there has been further consolidation and rearrangement of the idea units during the composition of the final EasyEnglish text. The identification and orderly display of the idea units is an essential intermediate step in the translation process. This step allows the writer to see clearly how to approach the translation in order to provide a logical flow of text that is as clear as possible.

# 2.4 Translation Techniques

We also use well-established translation techniques developed over many years by Bible translators. These aim to preserve the meaning of the text accurately and to ensure that the meaning is clear to the readers, who may be from a very different cultural setting from that of the original Biblical text. Barnwell (1999) gives the three most important qualities of a good translation as accuracy, clarity and naturalness. She says "Translation is re-telling, as exactly as possible, the meaning of the original message in a way that is natural in the language into which the translation is being made." The Biblical translation process demands a meaning-based (as opposed to word-based) approach. Larson (1984) has explored this aspect in depth. Such an approach, for example, will ensure that culturally-specific metaphors are translated appropriately.

A special problem, significant for Bible translators but which any translator may encounter, is implicit information. This is data the reader needs to know in order to understand the text, but which is not explicitly stated because the reader is assumed to be within the socio-cultural locus of the writer, and to know these facts already. Readers outside this original cultural context may well need this data to be overtly stated. For example, in 2 Samuel 15.32 we read of David's trusted friend Hushai meeting him "with his clothes torn and with earth on his head" without further explanation. Our knowledge of the culture informs us that Hushai had not in fact met with an accident, but was demonstrating his deep mourning and sadness in the traditional way of a 10th-century BC Palestinian. Our readers need us to explain this to them. This important facet of translation has received much attention in Biblical translation (Deibler, 1993; Deibler, 1999).

# 2.5 Comparison with Existing Easy-Reading Bible Translations

Other easy-reading Bible translations exist. Examples are the Bible in Basic English (BBE) edited by S.H. Hooke (1949), the Easy-to-Read version (ERV) from the World Bible Translation Centre (2000), and the New International Readers' Version (NIrV) (International Bible Society, 2001). As we might expect, there are many points of congruence between these translations and our own EasyEnglish translations. But there are also very significant differences, because of a difference in the intended audience. These three translations are aimed primarily at native English speakers (or sub-sets within this group; the ERV and the NIrV are particularly suitable for children). In contrast, our own translation is designed with second-language English speakers particularly in mind. The style of these three translations thus contrasts with that of our own in significant ways. For example:

- 1. All three use the passive voice.
- 2. The NIrV and the BBE make greater use of abstract nouns (e.g. '*faith*', '*mercy*', '*grace*').
- 3. All, to varying degrees, use more complex sentence structures (especially BBE).
- 4. All have a wider vocabulary, if one takes into account both the total corpus of vocabulary and the range of meanings used.

# 3 Production Aids

# 3.1 Printed aids

These three primary tools are currently used by our production team:

- 1. The 'EasyEnglish Procedure Manual' sets out translation techniques, rules for grammatical structure, and house style.
- 2. The 'EasyEnglish Lexicon' lists the vocabulary and the meanings allowed and not allowed at different EasyEnglish levels.
- 3. The 'Biblical Glossary' lists EasyEnglish definitions for words not allowed.

## 3.2 Computer tools

We use a computer program to check that the text conforms to some of the EasyEnglish rules:

- 1. **Readability index** (using MicroSoft Word's built-in Flesch Reading Ease Index). We aim for a readability level of 85.
- 2. Word macro add-on (produced by John Williams of the EasyEnglish team). This highlights (1) all non-permitted words, and (2) permitted terms that have non-permitted meanings commonly used in standard English.

## **4** Future Developments

#### 4.1 Software developments

The following developments are being considered:

- 1. Drop-down menus of EasyEnglish alternatives for non-allowed words.
- 2. Dynamic on-screen readability scoring by document, paragraph, or sentence.

#### 4.2 EasyWords Thesaurus

The aim of this tool is to help writers select vocabulary and phraseology at the appropriate level. It would do this in three ways:

- 1. Helping writers select EasyEnglish equivalents for words or phrases not allowed. If an EasyEnglish writer has struggled to find a way of expressing a complex idea in EasyEnglish, we need to share that writer's solution with other writers.
- 2. Suggesting alternatives to permitted words or phrases. It is easy to forget allowable (and potentially better) alternatives when writing especially at Level B, with its larger vocabulary.
- 3. Listing groups of related permitted terms (e.g. nautical, cookery and agricultural terms). This provides a 'palette' of terms to choose from when doing creative EasyEnglish writing (rather than translation).

# 4.3 EasyOtherLanguages (e.g. EasyFrench, EasySwahili, EasyPortuguese)

We would like to make our materials available in 'restricted' forms of other languages of wider communication. Both manual and mechanised translation are under consideration. Manual translation of EasyEnglish materials into other languages may well yield equally simple documents in the target language. However, commercial software may give good results, because the complexity of the input document has been greatly reduced through the EasyEnglish translation process.

Establishment of EasyOtherLanguage lexicons and grammars for manual translation, as we have done with EasyEnglish, is being considered. One problem (already mentioned) is the difficulty of obtaining corpora that use simple 'everyday' language. For example, in a Portuguese corpus we looked at (taken from a Brazilian intellectual newspaper) words like '*electronic'*, '*political'*, and '*Microsoft*' had higher frequency than appropriate to our audience; conversely words like '*baby*' and *'bread'* had a lower frequency than we would expect in everyday spoken and written material. How can we collect large corpora of transcribed spoken, popular forms of speech?

We also have to ask whether English-based readability scoring will work in other languages. These scores are based primarily on the length of paragraphs and sentences, and on the average number of syllables per word. As long as adjustments are made for these parameters especially the average number of syllables per word - such readability scores are probably universally applicable. In general, shorter words, sentences and paragraphs are likely to make a text easier for inexperienced readers, whatever the language.

# 5 Critique of EasyEnglish

## 5.1 Use of EasyEnglish Materials

Despite the difficulties of matching our vocabulary to our audience, we are confident that our Easy-English materials are much easier to understand than standard, unrestricted English. Circumstantial evidence of this is provided by the continuously increasing use of our material (gauged by the amount of our material downloaded).

## 5.2 Features of EasyEnglish

What specific features of EasyEnglish can we identify that might be the most significant contributors to the widespread use of our material? Perhaps the most important is the attention we give to making the text clear and meaningful to non-native English speakers spanning a wide diversity of mother tongues and from a great variety of cultural backgrounds. A somewhat similar approach is taken by such systems as the AECMA Simplified English (www.aecma.org). However, the AECMA system (as an example) is designed primarily for writing technical documentation, and thus serves what may be considered a narrower audience with perhaps a more restricted range of cultural diversity.

Our efforts to make our texts clear to such a diverse audience are implemented by applying established and sophisticated principles of Bible translation developed over many years for communication across cultural boundaries. The resultant high level (in our view) of cross-cultural transparency of our EasyEnglish texts may well be a particular encouragement to people to use our material.

#### 5.3 Field Testing and Feedback

We do, however, need to evaluate how well we are achieving our goals. There are two issues here:

- In many areas served by our material, it is culturally unacceptable to 'criticise the teacher'. Almost all our feedback is complimentary. This presents a big problem. How do we identify readers who feel comfortable about making negative critical comments, and how do we get feedback from them? One way may be to ask readers for ways in which the document could be improved.
- 2. Many people may feel satisfied with our material. But how many actually understand it as we intended? We need to pinpoint any of our material that is not easily or correctly understood by our audience. We also need to identify tools that can test comprehension.

We are beginning to explore comprehension testing with members of our target audiences to establish how well our EasyEnglish materials actually perform.

## 5.4 Criticisms of EasyEnglish

Furthermore, we recognise that there are valid criticisms of our approach. These include the following:

- 1. Our almost complete exclusion of passives can sometimes force us to find subjects for verbs when these are either not known or not important, or lead us into circumlocutions that do nothing to simplify the text.
- 2. EasyEnglish text can exhibit lack of flow or 'choppiness' - largely due to the restrictions on the number of clauses and on embedding. This can in turn upset the balance of emphasis in the text and hinder readers from distinguishing between background and foreground information.
- 3. Non-permitted terms often need to be replaced by phrases, and this can also introduce complexity and further disturb the balance of emphasis.

Greater flexibility in applying the EasyEnglish rules (for example, allowing more passives and relaxing the limitations on the number of clauses) may help the text to flow and restore proper emphasis. But this will then complicate the Easy-English grammatical rules, and consequently add to the writers' burden and demand more advanced writing skills of them. This in turn may lead to other, different, failures in the resultant EasyEnglish texts. We have to strike a balance between the needs of the reader and the demands on our writers, and this is not necessarily an easy task.

## 6 Conclusion

We believe that there are immense opportunities for the EasyEnglish system. It is estimated that between a quarter and a third of the world's population uses some English every day, and this is expected to rise to 50% by the year 2050 (Whitworth, 2002). This is fuelled by the rise of English as a medium of business and education. One index of this is the fact that a large majority (perhaps around 85%) of Websites are in English. Most speakers of English are not mother-tongue speakers, and many would benefit enormously from an English written specifically with them in mind.

Our major challenges are threefold: (1) developing a suitable EasyEnglish corpus; (2) adjusting our grammatical rules to allow greater flexibility whilst not unduly complicating instructions to our writers; and (3) obtaining feedback from our users. Overcoming these hurdles would help us to fine-tune the EasyEnglish system and help us to realise its full potential in communicating across cultural and social boundaries worldwide.

## Acknowledgements

Thanks are due to David Landin, Director of Wycliffe Associates (UK), for some material presented in this paper, and for his most helpful comments and suggestions.

# References

- Anon. 2002. Feedback. New Scientist, (7 December, 2002): 92.
- Barnwell, Katharine. 1999. *Bible translation: an introductory course in translation principles* Third edition (revised). SIL International, Dallas, Texas.
- Bennett, Karen M. Undated. *EasyEnglish training course* Wycliffe Associates (UK), Rhyl, United Kingdom.
- Bernth, A. 1997. EasyEnglish: a tool for improving document quality. In *Proceedings of the Fifth Conference on Applied Natural Language Processing*, pages 159-165.

- Deibler, Ellis. 1999. An index of implicit information in Acts-Revelation. SIL International, Dallas, Texas.
- Deibler, Ellis. 1993. An index of implicit information in the Gospels. SIL International, Dallas, Texas.
- Francis W.N. and Kucera H. 1982. *Frequency analysis* of English usage. Houghton Mifflin, Boston, USA.
- Hindmarsh R. 1980. *Cambridge English lexicon*. Cambridge University Press, Cambridge.
- Hooke S.H. 1949. *The Basic Bible, containing the Old and New Testaments in Basic English*. Cambridge University Press, Cambridge.
- International Bible Society. 2001. New Light Bible: a message for life. New International Readers Version. Hodder and Stoughton, London.
- Johannsson S. and Hofland K. 1989. *Frequency* analysis of English vocabulary and grammar: based on the LOB corpus. Clarendon Press, Oxford.
- Larson, Mildred. 1984. *Meaning-based translation: a guide to cross-language equivalence*. SIL International, Dallas, Texas.
- Thorndike E.L. and Lorge I. 1938. *A semantic count of English words.* Columbia University Press, New York.
- West, Michael. 1950. English Language Teaching. British Council, London. <u>as quoted</u> in Gauntlett J.O. 1966 'Teaching English as a Foreign Language' Macmillan and Co. Ltd., London, page 60.
- Whitworth, Damian. 2002. Attendez bien, mes amis. The English really don't need to bother with French. *The Times* (23 September 2002): 18.
- World Bible Translation Center. 2000. *Holy Bible: Easy-to-Read Version*. World Bible Translation Center, Fort Worth, Texas.