

Gained in translation ■

Science at the multilingual crossroads

“We believe that the privilege of being native English speakers comes with a responsibility to help EIL¹ colleagues with their English.”

John R. Benfield in *Chest* 2006;129:1728–30

For the modern-day researcher, English is the only way to gain wide visibility and actively participate in international scientific discourse. English is used by the most widely read and cited scientific journals. According to one study, the proportion of English publications in the area of chemistry increased from 54% in 1970 to 82% in 2000 [1]. The proportions of English papers originating in countries with official languages other than English also increased. For example, the share of English papers authored by French scientists increased from 16% in 1970 to 93% in 2000 [1].

Between 2003 and 2005, the *Journal of Thoracic and Cardiovascular Surgery* received 59% to 63% of its manuscript submissions from authors whose native language was not English, and the experience of the *Annals of Thoracic Surgery* was similar [2]. In both journals, the acceptance rates of manuscripts authored by non-native English speakers has been reported to be essentially the same as that for native speakers.

1 EIL, English as an international language

While most non-native English scientists do not appear to question the prominent role of English in today's world of science—for most of them, the use of English in scientific discourse was a given when they embarked on their careers—the requirement to publish in a language other than one's mother tongue does present yet another hurdle to getting one's research published. One way to overcome this is to increase the awareness of the Anglophone world of this extra effort. Yet another is to increase funding for science translation and editing.

Brian Budgell and his team have developed an impressive spectrum of methods and tools to assist non-native speakers of English in getting a handle on the linguistic challenges of science publishing. Last issue's translation section introduced one of their projects, i.e., the Springer online concordancer. Here's more on how it all started and where it's headed—the potential synergies between EMWA and the Centre for Biomedical and Health Linguistics are obvious and waiting to be exploited.

Gabriele Berghammer

gabi@the-text-clinic.com

References:

1. Sano H. The world's lingua franca of science. *English Today* 2002;72:4.
2. Benfield JR, Feak CB. How authors can cope with the burden of English as an international language. *Chest* 2006;129:1728-1730.



Centre for Biomedical and Health Linguistics: Helping non-native English speakers get a grip on the lingua franca of science

by Brian Budgell

How and why we got started

The Centre for Biomedical and Health Linguistics began approximately 6 years ago in Japan as an informal collaboration of friends who were all involved, one way or another, in helping Japanese students and colleagues deal with the challenges of biomedical English. The spectrum of our activities ran all the way from teaching biomedical terminology to first-year health sciences students to collaborating on biomedical research and co-authoring research reports. All of us were struck by the ineffectiveness of language education in Japan, and we wanted to do something about it, beyond the piecemeal help that we were providing to individual colleagues or small groups of students. Additionally, those of us who were working in undergraduate education saw considerable room for improvement in assessment processes. The upshot of this was that we began working together to create teaching and testing resources that addressed our common needs. Initial

efforts were based on our collective intuition of what constituted good biomedical English and good teaching practice. However, it was not long before we began to question the validity of our own methods. We wanted to solve real-life communications problems, not just add more noise to the confusion of language sites on the web and language textbooks.



www.bmhlinguistics.org

Helping non-native English speakers get a grip on the lingua franca of science

> Those of us who speak English as a first language may not fully appreciate the challenges facing those who come to English as a second language. Nor are the societal costs of language barriers well understood. Most biomedical research papers are now authored by writers whose first language is not English. These writers are under considerable pressure to publish in English to reach the widest possible audience through high impact journals. This practice produces a hollowing out of the biomedical knowledge base within their own language, since one would not normally publish the same material in both English and a second language. This then puts additional pressure on students and young researchers to study in English. Studying for a graduate degree, or conducting post-graduate research is hard enough in one's native language, let alone in a foreign language. Thus, our collective experience in Japan was that language issues were a significant barrier to education and international collaborations. Furthermore, the interval from completion of research to publication was substantially prolonged. This could harm authors by eroding the novelty of their work. It also meant that the translation of research findings into improvements in health care practices was delayed. The outcomes of such delays are real and meaningful – increased suffering and death. Hence, effective communication within the biomedical publishing field is not some kind of luxury; it is an imperative.

What we do and how we do it

Thus, to improve our own efficiency, the next step in our development was to look at the methods that corpus linguists use to characterise language. A corpus is an electronic collection of literature assembled with the intention of representing a body of language. For example, we chose the five most widely cited journals in midwifery and perinatal care to represent the language of this domain, rather than trying to collect every minor journal in the field. After formatting and archiving, all of our electronic corpora are then compared, using specialised software, to a reference collection of modern general English to identify words which occur statistically more often in the domain or discipline under investigation. We next filter out rare words or words which are not widely dispersed across the literature. From this we arrive at a list of keywords—over represented words—which are likely to have special meaning within or significance to the language under investigation. We can also automatically identify commonly recurring phrases employing keywords, and words which commonly occur in the company of these keywords.

“Those of us who speak English as a first language may not fully appreciate the challenges facing those who come to English as a second language. Nor are the societal costs of language barriers well understood.”

When we adopted this strategy, our thought was that with these standardised methods we could identify the language that users genuinely needed—the keywords and phrases, and the grammatical conventions which set biomedical language apart from general English. In short order, therefore, we found ourselves involved in three parallel activities: identifying the languages of biomedicine and health, developing learning resources and developing a valid assessment tool (the Test of English for bioMedical Purposes - TEbMP) to measure communicative competence. Over time, colleagues from Malaysia, China and Australia joined us, as they were facing similar challenges in their own countries. In Malaysia, most university entrants are very good speakers of English as a second language. However, the transition to a third language, biomedical language, often presents unexpected difficulties. China is a special case where a historical burden of isolation has impeded communication on many levels. In Australia, overseas students represent a substantial proportion of health sciences and biomedical students. In fact, they are very important to the financial stability of Australian universities and in some instances this has led to recruitment of overseas students without adequate consideration given to their language needs.

Developing and analysing corpora

In the early days of our collaboration, some of us were lucky enough to receive financial support from our universities, permitting us to develop and analyse a number of corpora—collections of literature representative of particular disciplines. Initially, we looked at the languages in three areas: nursing, public health and midwifery. The findings of these studies were incorporated into our teaching and testing resources. Additionally, the publication of these results probably helped us to obtain a larger grant which has funded further research and resource development. We now have an additional half-dozen corpora in various stages of analysis, a respectable web site on which to post our resources, and about a dozen peer-reviewed research publications and conference presentations. These are all ‘points on the board’ with our respective universities, but more importantly we can see the results of our efforts in the achievements of students and colleagues. We are therefore very grateful to the EMWA for the opportunity to engage a wider audience of stakeholders, some of whom we hope will become collaborators.

Developing learning resources

Medical writers, editors and translators might be interested in the actual word and phrase lists which have been generated from our various corpora. These are posted under our *Biomedical Language* directory and will give a good idea of the terms which are favoured by the different disciplines. Our online concordancer, also accessed via our

Helping non-native English speakers get a grip on the lingua franca of science

“We often encounter instances where a phrase or sentence is grammatically correct and accurately conveys the writer’s meaning, and yet it still doesn’t sound right to the completely fluent user of the language.”

Biomedical Language front page, permits users to search our corpora for authentic examples of word and phrase usage. This tool may help to identify preferred phrasings and will assist editors and reviewers in determining the ‘attestedness’, or prevalence in the authentic literature, of unconventional turns of phrase. Attestedness is an important issue for advanced language learners and editors. We often encounter instances where a phrase or sentence is grammatically correct and accurately conveys the writer’s meaning, and yet it still doesn’t sound right to the completely fluent user of the language. In these cases, being able to demonstrate to the writer that their phraseology is unconventional vindicates the advice of the editor and gives the writer confidence in the editorial process. In addition to our concordancer, other open-access tools are under development and will be posted to our site once testing is complete.

Under our *Learning Resources* directory are a number of short courses which were originally targeted towards non-native English speakers, both students and researchers in the health sciences. However, even accomplished writers and editors may find these pages illuminating. Regrettably, much of the popular advice that is offered to biomedical writers these days is based on conventions from general English—in some cases quite archaic English at that. By way of example, the avoidance of self-reference, especially ‘I’ and ‘we’, and the high prevalence of the passive voice are, given the culture of science, appropriate in biomedical writing and to be encouraged. This might not be so in general English, but in mastering any language, including biomedical language, we want to study it as it is, rather than as we might wish it were.

Developing TEbMP

The *Test of English for bioMedical Purposes (TEbMP)* is actually our oldest project, and is probably our most ambitious. Experience has shown that health sciences and biomedical students for whom English is a second language struggle with reading and writing regardless of performance on standardised English tests. Similarly, health professionals for whom English is a second language often encounter difficulties in communicating with their patients and colleagues, despite having met benchmarks for general English fluency. Clearly, popular tests for assessing competence in the languages of the health sciences and biomedicine are not making the grade.

As the name suggests, the Test of English for bioMedical Purposes is designed to assess competence in the authentic languages of biomedicine and health. Test questions are based on our corpora and are designed to evaluate

competency in the domains of reading comprehension, writing, listening and speaking. At this point, test questions have been piloted and validated with student cohorts in Australia, Canada, China, Japan and Malaysia. Our hope and intention is that, for entrance into biomedical and health sciences programmes, this test will replace or supplement tests of general English. The test may also find application in assessing the language abilities of health practitioners moving into an English speaking environment.

How we became involved with Springer



The ethic of studying authentic biomedical language and then emulating the most effective conventions is gaining currency. We are most pleased that a number of biomedical publishers have approached us for guidance on the development of learning tools for their own stakeholders. Having worked previously with Springer, among the many editors of their *Encyclopedia of Neuroscience*, we have been able to make connections with their e-products division and collaborate or consult on interesting and enjoyable projects. Our colleagues at Springer have also been very generous in providing access to their archives for the purposes of our linguistics research, which has resulted in the publication of a highly useful biomedical concordancer produced by Springer—www.springerexemplar.com. This resource is based on the enormous archives of this publisher and is undergoing continual refinement.

For the future

The Centre for Biomedical and Health Linguistics has been a collaboration of like-minded volunteers since its inception. Our intention is to continue to seek financial and in-kind support so that we can conduct important linguistic research, make the findings and applications of research freely available, and develop open-access resources for all stakeholders with an interest in improving communicative competence within the domains of biomedicine and health. A significant challenge at this point in our development is rising above the background noise to be recognised as an exceptional and credible source of reliable information on the lingua franca of biomedical and health communications. To stay true to our objectives, we also require feedback from our stakeholders, and objective measures of the effectiveness of our efforts. All of this will take many hands, and we hope that some of our colleagues in the EMWA will enjoy working with our organisation.

Brian Budgell

Canadian Memorial Chiropractic College
Toronto, Canada
bs.budgell@gmail.com