SCIENTIFIC WRITING ASSISTANCE FOR NON-NATIVE SPEAKERS OF ENGLISH: SHIFTING RIGHT ON THE INTERACTIVITY SPECTRUM

Ray Cooke
Susan Birch-Becaas
Université Victor Segalen Bordeaux 2

ABSTRACT

This paper investigates the issue of how non-native speaking (NNS) scientists manage to get their work published despite the linguistic and procedural difficulties that they face, and in particular, looks at the means and tools available to do this. The authors pinpoint certain conditions that could be met in order to allow NNS scientists easier access to the “gate-kept” world of publishing in English. Finally, they describe how available tools to help such authors have progressively shifted right on the interactivity spectrum.

KEY WORDS: NNS scientists, publishing research, gate-kept access, online writing assistance.

RESUMEN

En el presente artículo se analizan los mecanismos que permiten a los científicos cuya lengua madre no es el inglés publicar sus trabajos de investigación a pesar de las dificultades lingüísticas y de procedimiento con las que suelen enfrentarse. En particular, se revisan los recursos y herramientas disponibles para lograr dicho objetivo y se señalan ciertas condiciones que deberían cumplirse para facilitar el acceso de estos investigadores al restringido mundo de las publicaciones en inglés. Por último, se ofrece una descripción de cómo el grado de interactividad en relación con el uso de las herramientas existentes ha ido aumentando progresivamente.

PALABRAS CLAVE: científicos hablantes de inglés como lengua internacional, publicación de trabajos de investigación, acceso restringido, ayuda online para autores.

1. INTRODUCTION

English has long been the dominant language in the world of research and at present, almost all scientific research articles are written in English. Both novice and more experienced researchers the world over have to write in English for their work to be published. While publishing in English is the only way for these researchers to establish themselves as expert members of their discourse community,
Their results need to be made available to the widest readership possible. Hyland sums up the situation as follows: “countless students and academics around the world must now gain fluency in the conventions of English language academic discourse to understand their disciplines, to establish their careers and to successfully navigate their learning” (Hyland, *English* 24). Those hardest hit by what Benfield and Peack term the “English language burden” are of course the researchers for whom English is not their native language. The problems faced by these researchers have been well documented and summarised by Flowerdew (“Writing”) and Uzuner. Non-native speaking (NNS) authors are at an obvious disadvantage as they have to structure their papers according to cultural norms and conventions which are not their own and, of course, the time taken to draft an article is longer. The task is even more difficult for those researchers who live and work “on the periphery” or “off-network” (Belcher) as they may not have had any training in scientific writing (Salager-Meyer), they may not have access to assistance from native speakers (Li and Flowerdew) or they may not have the funding necessary to collaborate with an author’s editor (Flowerdew, “Scholarly”).

### 2. PREVIOUS AND RECENT TRENDS

A number of recommendations and suggestions have been made recently in order to assist NNS researchers with the publication of their research. Salager-Meyer stresses the importance of open-access to journals for researchers in developing countries and the first International PRISEAL Conference (Publishing and Presenting Research Internationally: Issues for Speakers of English as an Additional Language) at the Universidad de La Laguna, Tenerife in 2007 has also led to a statement on equitable access to the international academy. Salager-Meyer also underlines the need to develop initiatives whereby experienced researchers would help authors in the developing world to publish by providing editorial assistance and acting as mentors. The issue has also received attention from an editor-in-chief of a medical journal. Charlton suggests that NNS scientists should be permitted to construct papers using a substantial number of direct quotations from the already-published scientific literature. He argues that the result should be better than insisting that scientists of poor linguistic competence be compelled to re-phrase and recombine concepts which have already been well-expressed elsewhere in the scientific literature. Of course, ideally NNS researchers would have their manuscripts re-read by both an experienced peer and a language specialist as advocated by Flowerdew (“Scholarly”) and Benfield, but this is not always feasible especially in developing countries.

There is, therefore, a need for tools and materials to be developed to help NNS authors with writing up research. Previous research has provided valuable input for writing courses. Genre analysis studies have provided model moves which NNS researchers can follow (Swales; Hopkins and Dudley-Evans) and more recently other research genres such as letters to the editor have been studied (Magnet and Carnet). Other studies have focused on prominent language features of the
research article, and contrastive studies have also helped to illustrate how different cultures may organise their research according to different norms, with recent studies on Italian writers (Giannoni 2008), Sudanese authors (El Malik and Nesi) and French authors (Rowley-Jolivet and Carter-Thomas). However, with the exception of textbooks such as those by Swales and Feak \textit{(Academic Writing, English in Today's Research World)} and Weissberg and Buker, there are few pedagogical materials available for graduate students and researchers. They do not have the time to study the comprehensive style manuals available, and the “instructions to authors” provided by journals almost never give any readily transposable advice about the grammatical and strategic issues involved in drafting scientific English.

A recent trend has been the move towards the study of large corpora of academic discourse. Concordancing tools can be used to identify key words, recurrent grammatical structures, collocations and to draw up word lists. However, Tribble has warned against the use of vast corpora such as those used by lexicographers and descriptive linguists which, “while it may be valid for these communities, does not necessarily meet the needs of teachers and learners in EAP programs.” (Tribble 132). He advocates constructing a corpus with the type of texts that learners need to write: “the generalisations which they are able to make can be incorporated into their own written production” (ibid: 146). Flowerdew too has called for corpora of texts from the students’ own environment that reflect their writing tasks “exploiting small ‘localised’ expert corpora for pedagogic purposes” (Flowerdew, \textit{Corpus-Based} 134). The approach then should be to combine the bottom-up approach with a top-down analysis of the macro-structure of texts as in Flowerdew’s study of problem-solution texts (\textit{Corpus-Based Analyses}) or Hyland’s study (“Activity”) of reporting practices. Charles too argues that lexico-grammatical patterns revealed in corpora need to be associated with their functions in the discourse, “the use of corpus consultation to explore a series of lexico-grammatical patterns does not in itself add up to a coherent set of teaching materials” (Charles, “Reconciling” 290).

It would therefore seem that smaller corpora that can be examined manually and processed didactically may reveal interesting results and applications for the learner and EAP instructor. This raises the question as to what type of model NNS learners could follow.

3. ACCEPTABILITY AND GAINING ACCEPTANCE

This issue of how to define what is an acceptable level of English for publication has been raised both by Salager-Meyer and Flowerdew (“Scholarly”). NNS researchers should not be expected to produce native-like discourse but articles which are sufficiently clearly written to move forward to the review process. At the same time, editors should be encouraged to tolerate more “deviation” and “linguistic peculiarities” (Salager-Meyer 126). Indeed as Salager-Meyer points out, the distinction between novice and expert writing may be more relevant than that between native-speaker and non-native speaker productions. So far the tendency has been to study native-speaker corpora but it is perhaps as realistic for NNS authors
to imitate their successful NNS peers who have succeeded in being published. We would argue that NNS texts which have been re-read by an author’s editor and accepted for publication can be used as a realistic goal for novice researchers to achieve.

While it is one thing to write correct English, it is another to negotiate one’s way through the submission and acceptance process. Pitfalls abound. On the authors’ side—and to name but a few—there is: not adopting the journal’s “house style” or respecting the Instructions to Authors contained in every journal; lacking concision; not using linguistic devices likely to highlight the salient points of one’s research; not paying attention to the form of certain documents accompanying the main item submitted. The situation would be difficult enough if it were not for the fact that, on the journals’ side, matters have also become stricter: as many as five reviewers to satisfy; further experiments or statistical tests required; bibliographical references to be added; and, last but not least, level of English to be improved. Indeed, perseverance and toeing the line tended ultimately to prove fruitful and lead to publication in a study of NNS submissions to *English for Specific Purposes* (Belcher).

While these factors are perhaps the visible barriers to publishing in English, others lie in wait. If sufficient financial resources were available to outsource a problematic manuscript to a reputable scientific sub-editing agency, it is likely that it would gain not only linguistically but also structurally, since both form and function would probably be commented upon and corrected. However, being off-network probably means lacking the financial resources to outsource, and even lacking the human resources (proficient thesis supervisor, reliable peers, etc.) that could offset the problem. An even more unquantifiable but potentially damaging factor in the long term could also be the attribution of an off-network thesis supervisor at the outset of one’s doctoral studies. These are just some of the obstacles faced by NNS researchers in the quest to pass by the gatekeepers of English-language journals.

Indeed, if one attempts to summarize the present situation regarding publishing in English-language journals, a particular word springs to mind: pressure. There can be no doubt that publishing in English-language journals is not getting any easier, especially for NNS researchers. Years of personal observation lead to the conclusion that the pressure under which NNS researchers find themselves has steadily increased over the years. First, every country in the world maintaining a policy of funding state research lays down stringent requirements that researchers must meet with regard to the volume and type of publications that they are called upon to contribute to the accrual of center knowledge, i.e. the knowledge on which the mainstream scientific community is founded. France is also taking that avenue at present with the result that pressures are building in the research community. Indeed, a French government think-tank recently projected a 10% increase in the

---

1 25 years of experience helping researchers in a life sciences university to get their research published in English.
number of doctoral students in France between 2006 and 2015, the figure rising to 15% for the hard sciences and the health sector. Second, there is a growing tendency for the number of reviewers of submitted papers to be increasing. The journals with the greatest impact factor tend to have the largest number for reasons of scientific credibility. However, meeting with the approval of 4 or even 5 reviewers may not be easy, especially for NNS researchers likely to commit slight but annoying mistakes in English. This can only add to the pressure on their shoulders. The corollary is that a manuscript should be well written from the outset... or risk outright rejection, irrespective of the value of the science that it may contain but of which the editor-in-chief may perhaps be unaware. Furthermore, there is evidence that, with some rare exceptions, editors no longer can afford to take the time that they may previously have taken to rectify the faulty language of NNS researchers. This is understandable given the sheer volume of work falling on their shoulders in an era where they may have to deal simultaneously with both paper and electronic publications. For a number of reasons, therefore, researchers and editors alike are coming under increasing pressure in an ever more competitive research world.

Appropriate language training well upstream in a young researcher’s career might considerably alleviate the problem. In France, however, practices differ widely for reasons beyond the scope of this paper, some doctoral students having received, and continuing to receive, fully appropriate training in academic writing, with others receiving none at all. If the institution fails at this level, the problem of disseminating one’s research becomes an even more uphill one. In this context, writing up research can be a solitary frustrating experience, since fledgling researchers are expected to play a game whose rules may be known to them, but for which they may be ill-equipped in terms of skills. Who, or rather what, can they turn to?

To answer this question, it is important first to consider in what circumstances researchers write up their research. Anyone who has set foot in a research laboratory will know that researchers’ workstations tend to be rather cramped places with papers, notes and articles set around a computer screen and keyboard. In fact, the question of how researchers write up research probably involves as many ergonomic issues as it does linguistic factors. In a recent discussion, one researcher mentioned to the authors of this paper that he strews around his table top all the downloaded pdf versions of articles relevant to his own subject, sets dictionaries next to them, and then tries to “imitate the style of well-known authors” in his domain. Such personal resourcefulness merits research in its own right and might reveal some interesting process-driven insights for EAP practitioners. For example,

---


4 This could consist in well-planned ESP courses from Master’s level up.
in a study of 13 Japanese researchers (Okamura), only 5 paid direct attention to mastering English speakers’ language use, the others simply resorting to piecemeal observation and attempted imitation of what they read day in and day out in their profession. Whether these findings can be extrapolated to other national groups remains to be established. In any case, the tools that linguists make available to researchers to meet their needs have until now been linguistically oriented, and have by and large paid less attention to the ergonomic and cognitive needs of those for whom they are intended.

4. INTERACTIVITY AND EXISTING APPROACHES

We shall now turn our attention to the notion of interactivity. The term can have many meanings but it is human-computer interaction that the following discussion focuses on. Indeed, we are all aware of how an inadequate online user interface can be so damaging for the browsing and acquisition of otherwise well-founded information. Today’s busy researchers need a special kind of help available at their fingertips. Any other presentation format is likely to disinterest them, at least during the phase when they are actually drafting their manuscripts. In a recent study exploring the bibliographic and documentary information (BDI)-seeking behaviour of high-level French neuroscientists, the subjects were asked to specify exactly what they expect from a “good” BDI resource. Characteristics cited as being necessary were: fast, enabling users to master it on their own, exhaustive, interface easy to use, user-friendly and comfortable (Vibert et al.). We shall therefore examine how materials and procedures used until now can be viewed in the light of these requirements. Of course, modalities designed years ago cannot be fairly compared with recent advances, given the infinitely greater calculating power that modern computers offer to programme designers. The sole intention of the discussion below is therefore to demonstrate to what extent such digital materials meet the requirements of NNS scientists.

The simplest approach until now has been to take the contents of a paper book dealing with scientific writing and simply convert them into digital text. This solution, giving users the impression that they are doing nothing other than reading a book on screen, goes no further than the initial paper format. The problem for users is of a cognitive nature. While the search for relevant and reusable information involves the familiar action of turning pages, the extraction and transfer of appropriate structures and occurrences places excessive effort on the cognitive resources of the busy user. In this approach, the macro and micro writing skills required to produce acceptable text are not made available in a readily usable, model-based display. Indeed, it might even be that NNS scientists are likely to glean as much, or even more, pertinent information about scientific text construction simply by reading the scientific articles and other writings that they consult on a regular basis for their work. This hypothesis could be tested in an experiment with two groups matched for age, level and research domain. In any respect, this sort of approach lies well to the left on a progressive interactivity continuum.
Another approach has been to construct what may loosely be termed as “online phrasebooks” that are organized on a functional or notional basis. For example, the contents may include a dozen or so phrases or sentences relating to the notion of working at the laboratory bench. Such online phrasebooks may fall somewhat short of the needs of NNS scientists in their search for the keys to the structural discourse of scientific writing that they need to make them accepted members of their scientific community. The problem is that this sort of approach provides little real assistance since the phrases or sentences they contain remain fixed in stone, making it virtually impossible for the user to appropriate them for his/her own communicative usage. Hence, the degree of interactivity remains low because, once a particular type of occurrence has been consulted and (presumably) modified according to the user’s requirements of the moment, the database no longer serves any immediate purpose.

Another modality has been to approach a single aspect of the issue from the perspective of the course requirements of university students. Many universities throughout the world propose such online assistance in the form of “writing workshops”. Here, the aim is to make information available on some of the genres that budding writers such as NNS scientists need to master, for example, cover letters for résumés. On the other hand, such sites are also likely to contain advice on other purely university-based genres such as dissertation openers and closers, that is, the sort of advice that NNS scientists are less likely to need. Furthermore, a drawback of some writing workshops is that the information contained by such sites may be accrued by faculty members adding files to an existing text corpus, and that at no point in the accrual process does any single contributor go beyond this stage and programme the didactic processing of the corpus necessary for making the database truly interactive in order to allow heuristic learning. Even so, while not being exclusively designed for the NNS user, some writing workshops go well beyond the requirements of the students of their own universities and offer a treasure of procedural information on how a wide range of genres are to be approached, coupling this information with advice on grammar and style guides. Writing workshops may propose interactive feedback for the students of their own universities, and as such represent a step right on the interactive continuum.

Yet another approach, based on a long-time unquestioned tenet of teaching English as a second language, has been to establish the Anglo-Saxon model of English as the target model to be learned and imitated. Online tools using this approach implicitly posit, therefore, that the writings of Anglo-Saxon authors are the only models that others should attempt to imitate. Beyond the debate regarding the hegemony of one model over any others that lies outside the scope of the present paper, it is of note that a significant proportion of the research articles published in the world today come from non Anglo-Saxon countries. While their authors may

5 For example, see The Purdue OWL, <http://owl.english.purdue.edu/>, and the University of Toronto writing lab, <http://www.utoronto.ca/writing/advise.html>.
have sojourned as a PhD student in an Anglo-Saxon country, they likely do not possess oral communication skills fully equivalent to those of a native speaker. On the other hand, their written scientific English, although recognizably not of native-speaker origin, may reach a sufficient level of accuracy to be accepted for publication in the international journals of their specialty. Moreover, due to interlanguage, the phenomenon whereby the members of one language group such as the French are likely to commit similar errors when they write or speak a foreign language such as English owing to interpenetration between the two languages, it paradoxically becomes profitable in learning terms to take NNS productions and to transform them didactically into learning material for particular NNS target groups such as NNS scientists. It is this hitherto largely disregarded approach that the tool described below adopts by highlighting the errors committed by Francophone scientists and transforming them into learning input.

The tool attempts to solve the above-mentioned methodological problems since it allows users to visualize the salient structural and discursive features of written scientific English that, when copied and pasted into a new word-processing file, will make the new manuscript more likely meet the requirements of acceptability for publication, providing the scientific content itself is also of acceptable quality. The tool is constantly active on the screen and users interact regularly with it to draft their own manuscripts. In this regard, it constitutes a further step right on the interactivity continuum.

5. TOWARDS AN INTERACTIVE WEB-BASED WRITING TOOL

In this section, we briefly describe the TYOS® (Type Your Own Script) project developed in Bordeaux and how it is used. The objective of TYOS® is to process a small corpus of texts covering several disciplines (medicine, biology, biochemistry, wine science, dentistry, psychology, pharmacology) and encompassing the various genres written by NNS researchers (articles, abstracts, case reports, letters, replies to reviewers, responses as reviewer). The corpus comprises the corrected first drafts of Francophone researchers, not their finally published texts. The rationale and construction of TYOS® have been described elsewhere (Cooke and Birch-Becaas). We believe that this small corpus of “acceptable English” —acceptable in that the drafts were subsequently accepted for publication after further in-house editing from the sub-editors of the journals to which the drafts were submitted— can be used to illustrate how scientific writing functions. The typical language errors of the first drafts can be exploited pedagogically and compared with the text editors’ corrections and reformulations. The learner’s attention can then be drawn to typical features such as frequently used expressions, verb forms, link words and

6 TYOS (Type Your Own Script), <www.tyos.org>.
Let us imagine a real case in point. Doctor Sanchez would like to write an article in which he will report on mesothelioma in workers occupationally exposed to asbestos. He is very familiar with the “building bricks” of his specialty, i.e. the lexical items that cover his field of study. More unfamiliar to him on a structural level, however, will be the “cement” to put the bricks together. Furthermore, he may lack the procedural generic knowledge of why and how sentences follow on from each other, notably in the introduction and discussion of the article he intends to write. This is because, like many researchers worldwide, he may somehow have missed out on acquiring this knowledge during his education, even though some of this knowledge is acquired intuitively by regularly reading the scientific literature. Faced with the task at hand, Doctor Sanchez may spread the major papers on the subject across his desk and settle down for the hours of hard graft that he knows he faces, as he highlights sentences or features in the papers that he thinks he could usefully exploit in his own writing. By consulting TYOS®, Dr. Sanchez would be able to structure his thoughts more clearly and presumably marshal more cognitive resources for the writing process.

The TYOS® corpus contains several full articles that have been processed didactically so that the user can obtain pertinent information on certain key aspects of writing that pose problems to NNS researchers: verb forms, discourse linkers, useful expressions, grammar points, vocabulary and discourse moves. The extract below shows one of the frames in a medical article visible to the user: useful expressions.

A deleterious impact of substance use on the course of psychotic disorders has been reported by several studies (1-4). However, other studies have suggested that subjects with a dual diagnosis may present with a less severe form of illness (5-8). These discrepancies may be due to methodological differences in the assessment of outcome, and to the heterogeneity of the samples mixing patients at different stages of the illness. Convergent findings demonstrate that subjects with early psychosis present with a high prevalence of substance abuse (9-13). As the first 5-10 years after the onset of psychosis predict the subsequent course of illness (14), it is of major interest to identify prognostic factors than can be modified by therapeutic interventions at this stage of illness. We previously reported in a sample of first-admitted subjects with psychosis that subjects with persisting substance misuse over the follow-up period are more likely to present with suicidal behavior and poorer medication adherence (13, 15). However, we did not explore the impact of substance and alcohol misuse on the other outcome dimensions. The aim of the present study was to further investigate the impact of substance and alcohol misuse on clinical and social outcome over a two-year follow-up after a first hospitalization for psychosis.

We thank Professor Hélène Verdoux, Department of Psychiatry, Bordeaux Teaching Hospital, for kindly supplying the draft version discussed in this article.
The grey highlighting shows Dr. Sanchez what “cement” he might be able to use from writing previously considered of sufficient quality to move forward to the sub-editing that most journals now undertake before publication. When he passes his cursor over each block of grey highlighting, he opens up a bubble containing an explanation why that phrase or occurrence could be re-utilized. For example, on opening the bubble concerning the occurrence “it is of major interest to…” (line 8), he sees the following explanation: “Says how important it is to do this. Therefore, author justifies his work”. If Dr. Sanchez then feels that he could reutilize this occurrence, he copies it to a word-processing file and builds the rest of the sentence with the “brick” language he is familiar with. For example, he could write: “it is of major interest to investigate the time lag to the onset of mesothelioma in workers exposed to asbestos”. Other aspects of scientific writing would also typically pose him problem. For example, the change of past tense from line 1 to line 7 might require an explanation. Bubbles would do this in the “verb forms” display. Finally, the macrostructure of the introduction is shown by a discourse analysis display based on Swales’ CARS model (Swales). It then becomes simple for Dr. Sanchez to identify the language of transition from one move to the next and to copy it to his own writing. Hence, by imitating a proven model, it becomes less likely that the finished product will contain errors of a structural or superficial nature that could be a hindrance to the reader.

6. STRENGTHS AND LIMITATIONS OF THE TOOL

This approach has limitations. First, marking-up texts is a time-consuming process (see Cooke and Birch-Becaas) so there is a problem of critical mass. Moreover, a standardized protocol must be respected by anyone who marks texts. To meet this challenge, the authors are currently working on a semi-automatic flagging system so that the corpus can be extended while maintaining its internal harmony and gaining greater power. Second, the tool still requires add-ons such as suggestions for pedagogical use and an extensive page of links. Third, there are still a number of residual typos and errors in version 1.0, although these are progressively being ironed out. On the other hand, the approach is transposable to other discourse types such as legal and business English, as well as to other languages. In this sense, TYOS® is not only stand-alone or classroom tool but, in the wider sense, is a generic methodology that can be applied in other settings.

The next step will be to transform TYOS® into a full productivity tool offering a wide range of discourse and genre types, coupled with some degree of automation and a set of online user-oriented services. Indeed, tools that will follow will no doubt contain even more in-built interactivity and will represent a further shift right on the interactivity spectrum.
WORKS CITED


—— *Corpus-Based Analyses of the Problem-Solution Pattern: A Phraseological Approach*. Amsterdam: John Benjamins, 2008.


