

Interdisciplinary perspectives on abstracts for information retrieval

Soon Keng Chan and Schubert Foo

Nanyang Technological University, Singapore

Abstract

The paper examines the abstract genre from the perspectives of English for Specific Purposes (ESP) practitioners and information professionals. It aims to determine specific interdisciplinary interests in the abstract, and to explore areas of collaboration in terms of research and pedagogical practices. A focus group (FG) comprising information professionals from the Division of Information Studies, Nanyang Technological University, Singapore, convened for a discussion on the subject of abstracts and abstracting. Two major issues that have significant implications for ESP practices emerged during the discussion. While differences in terms of approach to and objectives of the abstract genre are apparent between information professionals and language professionals, the demands for specific cognitive processes involved in abstracting proved to be similar. This area of similarity provides grounds for awareness raising and collaboration between the two disciplines. While ESP practitioners need to consider adding the dimension of information science to the rhetorical and linguistic scaffolding that they have been providing to novice-writers, information professionals can contribute useful insights about the qualities of abstracts that have the greatest impact in meeting the end-users' needs in information search.

Palabras clave: abstract, genre, English for Specific Purposes (ESP), information science, academic writing

Resumen

En este artículo se examina el género del resumen (*abstract*) desde la perspectiva de los profesionales del Inglés para Fines Específicos (IFE) y de la información. Se intenta determinar los objetivos interdisciplinarios específicos en el resumen y explorar áreas de colaboración en la investigación y en la práctica pedagógica. Un grupo de investigación de la División de Estudios de la Información, de la Universidad Tecnológica de Nanyang, en Singapore, se estableció para analizar el tema de los resúmenes (*abstracts*) y de los sistemas de resumir (*abstracting*). Durante los debates, dos temas importantes surgieron con implicación significativa en la práctica del IFE. Si bien las diferencias son aparentes en cuanto al enfoque y a los objetivos del género de los resúmenes entre los profesionales de la información y los del lenguaje, se demostró que las exigencias de unos procesos cognitivos específicos implicados en el proceso de resumir (*abstracting*) eran las mismas. Esta semejanza

proporciona la base para la concienciación y colaboración entre las dos disciplinas. Mientras los profesionales del IFE necesitan considerar añadir la dimensión de la ciencia de la información al andamiaje retórico y lingüístico que han proporcionado a los escritores noveles, los profesionales de la información pueden contribuir con percepciones útiles sobre la calidad de los resúmenes para hacer frente a las necesidades del usuario en la búsqueda de información.

Key words: resumen, género, Inglés para Fines Específicos (IFE), ciencia de la información, escritura académica

Introducción

The focus of English for Specific Purposes (ESP) has evolved over the last twenty to thirty years. ESP teaching began by adopting a lexico-grammatical pedagogy (Swales, 1985) for teaching science and technical subjects in English. The focus then moved to genre issues in the 80s and 90s (Swales, 1990). Currently, it is widely accepted that genre interests have become interdisciplinary. Genres are no longer considered static in form and structures, but are shaped by the contexts within which they operate (Freedman, 1999). Consequently, ESP teaching and research interests have also become multidisciplinary to provide for a more holistic approach. Interest in abstracts by ESP practitioners has also followed similar directions, even though this genre has not been as extensively studied as the research article (RA). Graetz (1985) was among the early ESP practitioners to study the abstract. She followed a learner-centered grammar approach to investigate abstracts from various publications. Later research interests began to emulate Swales' (1990) groundbreaking work on the "Introduction" sections of RAs. This trend extended to studies of the abstract as a genre in various disciplines, for example, in the medical field (Salager-Meyer, 1990; Hartley et al., 1996; Huckin, 2001). The main objectives of these genre-based studies have been on identifying and assessing language conventions and rhetorical structures of abstracts in various disciplines, and on drawing implications for improving classroom practice and helping second language learners.

While several ESP practitioners have acknowledged the growing significance of abstracts in the current electronic age of information retrieval (Posteguillo, 1999; Huckin, 2001) few, except for Posteguillo (2002), have expanded on how to incorporate this significance to teaching the genre. Thus the gap between what

information experts expect of abstracts and what ESP teachers practise in abstract pedagogy remains. Few ESP practitioners have worked with information professionals to establish a better understanding about information systems, or to identify the qualities of abstracts that would help end-users in information retrieval. The current study would like to suggest that establishing this link would further enhance ESP interests in the abstract. Moreover, the impact of technology information retrieval has further prompted this need to look at the abstract from a combined viewpoint of the two disciplines.

The paper proceeds with a review of research interests in abstracts by ESP practitioners and information professionals over a period of about thirty years (1960s to present). The results indicate differences as well as common grounds between them. Next, a focus group (FG) discussion among seven information professionals working in the Division of Information Studies at the Nanyang Technological University, Singapore, was conducted. The results of the study have implications for interdisciplinary collaborations and pedagogical improvements.

ESP practitioners' interests in the abstract

Since the 1960s, many English language teachers in tertiary institutions became instructors of non-native English graduate students whose poor command of the language prevented them from writing acceptably in their new disciplines. These teachers who became known as ESP practitioners based their pedagogy on the principle that English should be taught not as a grammar-based subject of the past but as a means for communication. Many approaches were developed for this purpose. The most popular practice was the genre approach developed by Swales (1990). Swales had defined the genre as a communicative event that has a common purpose and a mutual aim shared with other members of the same discipline or discourse community. A genre has structure and operates within standardised communicative constraints, and those who practise it frequently and professionally will have an overt knowledge of its rhetorical features. For example, research writers who publish their work in English are guided by conventions of the RA in that their papers must abide by the IMRD structure (Introduction, Method, Results and Discussion or Conclusion), and be expressed in specific language conventions acceptable to their discourse communities. This prompted many ESP researchers (Marshall, 1991; Gosden, 1995; Mustafa, 1995; Sionis, 1995; Paltridge, 1997) to recommend that academic writing instruction should

focus on teaching genres of specific disciplines. Besides that, much research was based on analysing the RA or specific sections of it. For example, Swales studied the Introduction section (1981, 1984, and 1990), and Dudley-Evans (1986) analysed the Discussion section of MSc dissertations. Nwogu (1997) studied RAs of medical articles, and Posteguillo (1999) examined RAs of computer science.

Similarly, interest in the abstract genre was also developed. Early writers who examined abstracts included Huckin and Olsen (1983) who applied Swales' genre analysis approach and examined language conventions of the genre. In the classroom, ESP teachers used abstracts as tools for their writing courses (Davis, 1991). The abstract is also viewed as a promissory preview (Swales, 1990; Salbiah, 2000) for conferences. However, recent research interests are reflecting more specific and deeper concerns.

Salager-Meyer (1990) studied abstracts from medical journals because of frequent criticisms received about badly written medical abstracts that were uninformative, misleading, and lacking in internal structure. Results of her analyses revealed that many abstracts had no purpose statement and no conclusion, and there was a prevalence of illogical sequencing in move organisation. There were also flaws in paragraph structuring and overlapping semantic concepts straddled between paragraphs. These findings prompted Salager-Meyer to propose that novice-writers should be given good models to emulate.

Keogh (1994) examined the structural and stylistic features of a corpus of 48 abstracts written by scientists and engineers for their workplace publications. For example, he found that textbook guides recommended the use of active structures but his sample showed an almost even mix of active and passive sentences. While academic texts had stressed the importance of conclusions and recommendations Keogh's samples seldom included them. Keogh concluded that in the real world a lot of academic advice tends to be discarded.

Santos (1996) studied how abstracts could be characterised in terms of textual organization, and analysed other key features of this genre. He found a prevalent five-move model with submoves: Situating the research, presenting the research, describing the method, summarising results, and discussing results. Santos concluded that his resultant schematic pattern or model would provide pedagogical advantages to non-native novice-writers in reading and writing abstracts.

Posteguillo (1996) focused on how discipline variations affect the abstract genre. He did not find the consistent use of IMRD structure of Swales (1990) in computer science abstracts, although he found similar structures exist. Although Posteguillo did not fully agree that the IMRD structure is used in computer science abstracts he agreed that abstracts reproduce structures of source texts. Posteguillo concluded that these results have pedagogical implications.

Berkenkotter and Huckin (1995) interviewed seven scientists extensively to examine their RA reading habits. They found that these scientists practised scanning for interesting new information. For example, they would read the title first and then the abstract to 'size up' critical information and decide on whether to read the full document. Next, they selected a corpus of 350 journal articles to examine how research papers structure and information organisation have changed to accommodate present-day reading habits of scientists who have little time and who are faced with information overload. They found that titles and abstracts now tend to be more informative to cater to the needs of current readers. For example abstracts have become inherent parts of papers, longer, and more informative, particularly in providing results and conclusions of studies so that readers may find it unnecessary to read the full document. In their study, Berkenkotter and Huckin also observed how abstracts have become significant surrogate documents in information gathering. As a surrogate, the abstract provides an alternative to reading the full article because it contains abbreviations of main concepts of the paper, namely, purpose, method, results, and conclusions of the source article. However, from these studies, it appears that although ESP interest in abstracts has increased, the interest remains largely genre-related.

In a recent study, Huckin (2001) examined the qualities of authored abstracts that have been put online, to determine whether they captured the main concepts of their full articles. He selected a corpus of biomedical abstracts and used the keywords provided by professional indexers (who are information professionals), to match with the expressions and concepts presented in the abstracts. Results showed that these abstracts closely reflected the main concepts of the articles. His study provided evidence that indexers and information retrievers could rely on authored abstracts to contain the main concepts of their full articles. This is a progressive move towards interdisciplinary research with the Information Studies (IS) counterpart. However, the present authors believe that direct collaboration with IS professionals could be more significant, and that their active contributions will provide better insights to pedagogical designs.

IS professionals' interests in the abstract

During the 60s and 70s, advanced technologies in information transfer resulted in a proliferation of information on electronic systems (Lorenz, 1969) and users faced problems in having to filter out information in their systems. Consequently, information (IS) professionals like indexers, systems designers, librarians began to accept abstracts as useful surrogate tools for quick and effective information retrieval. The conciseness and precision of abstracts make them time saving alternatives for information updating, and users could treat them as mini-texts to preview full articles. In the 80s and 90s (Hills, 1983), further technological advancements in high-speed computerization and Internet facilitated even more widespread and expedient information circulation. Full text information could now be efficiently retrieved and free text searching is possible. Nevertheless, Pinto and Lancaster (1999) remained convinced of the surrogate role of abstracts. They explained how full texts often contain details that cloud efficient identification of significant and precise concepts whereas abstracts would filter out unnecessary details. Similarly, Fidel (1986) and later Chowdhury (1999) added that full-texts retrieval and multimedia information in digital forms have in fact increased the usefulness of abstracts in scholarly publications.

Consequently, information professionals in indexing and abstracting (Collision, 1971; Borko & Bernier, 1975; Rowley, 1982; Cleveland & Cleveland, 1983; Lancaster, 1991) paid great attention to producing guides on abstracting. To them successful abstracts are characterised by qualities of brevity, accuracy, and clarity, exhaustivity, and on whether they serve the information needs of users. Other IS professionals instituted standards on abstracting; the most significant being the American National Standard Institute (ANSI Z39.14-1979), and the International Standards (ISO 214:1976). Prolific publications on prescriptions for abstract writing belie the great concern of information professionals for regulating abstracts qualities and abstracting skills. Therefore, it is not surprising for IS research on abstracts to be focused on evaluating and testing the qualities and usefulness of abstracts to end-users seeking information on the systems.

A number of studies were conducted to assess the readability of abstracts using readability formulas, comprehension measures, or both. For example, Dronberger and Kowitz (1975) suggested that the measurement of readability could provide an assessment of one phase of an information system. They explored abstracts published in *Research in Education (RIE)* whose full documents were stored in

Educational Resources Information Center (ERIC). Results showed that the reading level of an abstract was significantly higher than that of its full document because of its concise style and condensed form. The study also showed that measuring readability of abstracts is a useful evaluation technique.

Tenopir and Jacso (1993) measured the qualities of abstracts based on the following criteria: language style and readability levels, the extent to which the ANSI standard is observed, and the exhaustivity or breadth of content coverage represented in abbreviated texts. They found that readability was lowered by use of passive voice, inclusion of too many prepositions, too many sentences per paragraph, too many words per sentence, and too many syllables per word. However, the exhaustivity quality could not be so easily tested because of individual user needs. In another study, Hartley and others (1996) proved that users obtained faster and more accurate information from structured abstracts than traditional ones, implying that the abstract genre should evolve to meet the needs of current users.

In more recent studies, Wheatley and Armstrong (1997) conducted a study on abstracts produced for online services. Briefly, both IS professionals compared these abstracts to traditional abstracts by exploring the use of keyword descriptors, and metadata (author, journal, date of publication and reader assessments) on the grounds that such data would facilitate more efficient and accurate retrieval. In addition to these suggestions for changes to the traditional abstract, there is a strong promotion for human abstraction to be replaced by automatic text summarisation in information retrieval (Kan et al., 2001). However, Pinto and Lancaster (1999) who based their study on judging the qualities of authored abstracts in terms of exhaustivity, accuracy, readability, cohesion, brevity, and cost concluded that these qualities remain valid in the process of knowledge discovery. They added the dimension of assessing abstracts from different perspectives, such as from the product end, the process end, or the user end. Overall, it appears that although the computer has enabled easy availability of full texts in electronic forms, and automatic text summarisation is a few clicks away, author-prepared abstracts from the human readers' point of view remain relevant and necessitate further research.

ESP vs. IS interests in the abstract: Motivation for FG discussion

The review above has indicated that from the ESP perspective, the main interest in abstracts is related to linguistic issues, rhetorical structures, and language conventions of

abstracts in various disciplines, whereas from IS perspective qualities of readability, brevity, accuracy and exhaustivity are major concerns. These differences may be explained in the following statements. ESP practitioners are teachers and their clients are usually non-native speakers of English learning to write acceptably for their respective discourse communities. Therefore, pedagogical concerns are the ESP practitioners' primary motivation for investigating the abstract genre. Underlying this genre approach is the belief that if novice-writers are explicitly taught the rules and conventions of a specific genre, they will internalise the rules and produce acceptable work for their target audience.

ESP Professionals	IS Professionals
Pedagogical tool for teaching summarization Writer-centered: Teach models Emphasis on text-based and corpus-based research ---more recently: Awareness of technological influences	Retrieval tool for information communication End-user/reader centered Emphasis on technology and influence on abstracts research ---more recently: Abstracts may be replaced in due time

Figure 1. Abstracts as seen by ESP and IS professionals.

This belief has influenced practitioners to teach their clients to write abstracts that mirror the rhetorical structures and language conventions prescribed by respective discourse communities, international standards, journals, or conferences. It has also influenced ESP researchers to explore, examine and analyse authentic texts from various disciplines in order to test and verify these rhetorical rules so that their clients may benefit from their findings and expand their writing experiences. Consequently, ESP research on the abstract is highly text-based, corpus driven, and writer centered. Few studies have focused on the users or readers of abstracts, a perspective that is the preoccupation of IS professionals. This pattern of teaching and research continued until lately when the electronic medium became all pervading, and writers like Huckin (2001) began to shift focus and extend ESP research interest to online information retrieval. A summary of ESP versus IS interests in abstracts is shown in Figure 1.

In contrast, IS professionals have long viewed the abstract as a communication tool for retrieving information. It functions as the screening device for making decisions to read the source article and as an expedient means of information transfer. With advancements in electronic information systems, author-prepared abstracts have not only remained as information communication tools but have increased in importance because of information overloading problems. As such, the focus should be naturally

on readers or end-users, and the pervasive question should be whether an abstract fulfils its designated functions and meets the needs of its users. Just as ESP practitioners were prompted to focus on pedagogical issues of teaching, IS professionals were prompted to ensure that abstracts meet the needs of users, thus accounting for heavy dependence on prescriptive rules and attention on assessing qualities. Moreover, while ESP researchers remained focused on the abstract text and its rhetorical qualities, IS professionals have moved with the rapid changes in technology. In IS discipline, the future of abstracts is ever fluid. For example, concurrent developments in the use of metadata, keyword descriptors and automatic summaries are threatening to replace abstracts in online retrieval.

However, on closer examination, despite these differences, the perspectives of the two disciplines are not mutually exclusive. On the contrary, there are several areas where their concerns overlap or mirror each other, particularly, in areas of language concerns and structure. For example, the IS focus on 'brevity' is a language issue of concise writing skills. The IS concern with 'accuracy' and 'exhaustivity' are language matters of comprehension, interpretation and summarisation skills. And, the IS interest in 'readability' requires language skills in content organization, cohesion, cohesiveness, and clarity in expression. Based on these parallels, the present study is interested in confirming the observations with IS professionals in a live discussion. The objective is to explore opportunities for collaboration, and in the process, perhaps, uncover further nuances, opinions, recommendations, and interests of IS professionals with regards to using abstracts as tools for information retrieval and dissemination.

The focus group

A focus group (FG) discussion was conducted with a group of seven information professionals, whom Dudley-Evans (1986) and later Bhatia (1993) referred to as 'specialist informants'. The seven IS professionals were academicians in the Division of Information Studies at the School of Communication and Information, Nanyang Technological University, Singapore. These participants are highly experienced professors who offer courses in library science, information organisation, management, retrieval, systems design, and systems programming at postgraduate levels. Two of the participants are also journal reviewers and editors. Thus a comprehensive representation of information professionals whose opinions and comments are highly reliable and respected was secured for the FG discussion.

The FG discussion was held in a dedicated meeting room specially designed and equipped with both audio and video recording facilities, and a one-way glass partition that enabled unobtrusive observation during the live discussion. The room could accommodate a maximum of 12 participants seated around a large rectangular table (2m by 2m) that had built-in microphones at intervals of about a meter each. Built-in cameras could capture the images of the participants from all four angles. The discussion was scheduled for 90 minutes, and the entire session was to be audio and video recorded. The discussion was based on the following statement:

Current technological advancements in information dissemination and retrieval are posing an additional challenge to English for Specific Purpose (ESP) practitioners who teach academic writing. Is the current ESP focus on teaching rhetorical structures and linguistic conventions of genres enough? Have ESP practitioners a role in helping novice-writers meet the needs of abstracts for the web environment?

Three days before the session, a set of nine questions (Appendix) was e-mailed to the participants as a general outline for the discussion. On the day of the meeting, the first author assumed the role of moderator, following Morgan's (1998) guidelines for this type of research set-ups, and briefed the participants on the objectives and procedures of the discussion.

Results and discussion

In the course of the discussion, some of the original questions became redundant or overlapped in content and aspect. In addition, unscheduled but pertinent aspects of the abstract were raised. The significant issues that resulted from the discussion are summarised in following sections below.

Typological distinctions of the abstract

The IS professionals began by questioning the type of abstract that should be included in the discussion, an interesting point that had not occurred to the researchers to be of primary consideration for the discussion. The group proceeded to compare the typology of abstracts. For example, indicative abstracts are aimed at the larger and more generic readership, and do not normally contain results and conclusions, but are descriptive in style. On the other hand, informative abstracts are aimed at more restrictive audiences, and results and conclusions are of major interest

because they are focused on the latest progress made in their disciplines. Informative abstracts are usually written by authors of scientific and technical source documents. Review and critical abstracts, unlike indicative or informative abstracts, include the abstractor's comments or evaluations of the source documents.

In addition to this typological consideration, the information professionals raised another question: "Are we comparing original authored abstracts, or professionally produced abstracts, or authored abstracts that required modifications by professionals, or automatic text extraction?" In their opinion the latter type should not be discussed because of its mechanical nature, and the agreement was to concentrate on only authored abstracts.

Another consideration raised was, the choice of abstract type also depended on factors such as purpose, subject discipline, database source, and policy of abstracting agency. It was important to consider the purpose of using abstracts because purpose determines product variations (Cremmins, 1996) as well as the environment in which they are found. For example, informative abstracts are preferred in science and technical writings because they focus on results and conclusions, and are published in scientific or technical journals; indicative abstracts are usually shorter and merely provide abbreviated descriptions of the original documents. Indicative abstracts are common in database journals which provide only abstracts with bibliographic details. Another distinction is discipline. Different disciplines affect content and structures of abstracts. For example, it is common for abstracts in medical articles to be structured under subheadings, a feature not practised in other disciplines. Abstracts in mathematical or chemical engineering discipline use specific "unique identifiers" which are highly specialised like mathematical symbols, and they are used to aid faster retrieval.

The source of abstract databases is also significant. For example, LISA (Library and Information Science Association) and ERIC (Educational Resources Information Center) agencies do not have specific guidelines for authored abstracts, and would accept abstracts based on their subject matter. On the other hand, organisations like IEEE Transactions (Institute of Electrical and Electronic Engineering journals) and ACM (Association for Computing Machinery) would have explicit guidelines for authors. Lastly, an abstracting agency that aims to use abstracts to reach a global audience may decide that the indicative type is more appropriate. On the other hand, an abstract may be used in a book review, and in this case the critical abstract is

preferred because it includes an evaluation of the source document. Apart from the agency's policy of abstract usage, there are also considerations of cost and length, which could affect the choice of abstract type.

The typological distinctions cited above confirm the information professionals' focus on user needs in information retrieval and in information systems management. In contrast, while IS professionals make such specific distinctions about the types of abstracts; this typological perspective has not received as much attention in ESP studies. Instead, ESP teachers are observed to be paying more attention to teaching informative abstracts, which are common in scientific and technical writing. This is probably because of their focus on teaching novice-writers the RA. Based on this observation, there are pedagogical implications for raising novice-writers' awareness to typological distinctions of abstracts. For example, if an introduction to abstract types and related typological distinguishing factors are included in the course on abstract writing, novice-writers would develop a better appreciation of the value of abstracts and aim to write to cater to information retrievers and information systems.

Effects of technology on abstracts and abstract writing

Another issue discussed by the FG participants was whether abstracts that are increasingly put onto electronic information systems such as Internet, database collections and digital libraries have to be modified for the new environment? The major viewpoints are discussed below.

The participants were of the opinion that changes in environment are not responsible for affecting qualities of abstracts. For example, an authored abstract taken from a journal paper like IEEE Transactions or ACM and put into a digital library, or even on an electronic platform like Internet does not need to be rewritten or adapted. Besides, whether the abstract is used in print or Internet should not impact or affect its value as a retrieval tool. However, sometimes an original authored abstract could be modified by human assistance because of specific abstract or information agency objectives or policies, or because the original abstract is badly written. For example, the abstract agency may want to modify original abstracts so that they reach a wider audience on the Internet to cater to less technical readers. In this case the original informative abstract may be rewritten as an indicative abstract. Another environmental factor that could influence the abstract is the problem of

overwhelming information flooding the systems, thus compelling the length and format of the abstract to be modified. Finally, electronic abstracts that have been generated automatically by machines are different from the traditional ones, but these have no comparison with human authored abstracts; therefore, there is no need to be concerned with them.

After the discussion above, two suggestions closely related to technological intervention were raised. In the first, to facilitate cross-referencing to full texts, online abstracts could have hyperlinks in various parts of the document to cater to different needs of the user. For example, if the user is interested in only the results of the study a link could be made to the result section of the full document. All these links are only a click away, whereas in the traditional environment time and distance are constraints and such cross-referencing is tedious though not impossible. However, this utility is possible only if the abstract in question is accompanied by its source document and only if the abstract content covers all pertinent sections of its source. So far this idea has remained only a proposition.

The second suggestion was to include metadata as enhanced retrieval tools. These include using keywords, providing author information, number of words, and publication details as well as dates, assessment of its usefulness, and other utilities that would assist the user in retrieving information more efficiently. The question of whether abstract writers should learn to include such metadata to their abstracts is discussed in the next section.

Enhanced information retrieval (IR) tools in information systems

Arising from the various enhanced information retrieval (IR) tools on electronic systems that have been made possible by technological advancements the question was whether these utilities or facilities should also be taught to abstract authors so that they can incorporate them in writing.

The first utility discussed was the use of keywords. The IS professionals were of the opinion that including keywords to abstracts is important, provided they are carefully selected. For example, including the keyword 'transputers', which may not be present

in the original article, but is associated with the term 'parallel processing' would help to give the abstract a better chance of being retrieved. In order to know what keywords to select writers would need a good knowledge of the semantic web or thesaurus of the discipline in which they write. Including keywords from the semantic web would help the abstract to reach a wider audience. However, there is a drawback in using the facility. For example, over enthusiastic writers may 'abuse' the facility in order to ensure greater accessibility of their abstracts. Such an activity would cause 'spamming' problems and jam up information systems. One way to counter this problem is to ensure that only significant keywords are offered to help users obtain correct or relevant information from the whole gamut of information available on the systems.

However, the question was whether ESP practitioners should teach their learners to incorporate such utilities in their abstracts. The IS professionals were hesitant to recommend this responsibility, which they identified as an area in indexing. They explained that indexing is a specialised skill. It requires the indexer to first *read* the document technically, (that is, to read the article purely for the intention of selecting keywords), and then to use his or her expertise to pull out the right information from the various parts of the document for the abstract. Moreover, indexers do not depend on abstracts alone to select significant keywords. They usually have access to specific semantic webs, thesaurus or some unique identifiers that abstract writers may not have. However, they admitted that it is also true that in pioneering research and new frontier knowledge, keywords must first exist in the text of RAs before professional abstractors or indexers can become aware of them. In this respect, the abstract is very important because it is here that new terms are introduced and communicated by authors, before being introduced to information professionals as keywords.

The rating utility is also another facility available to assist abstract users during retrieval. By referring to the rating index abstract users could be assisted in determining the relevance and significance of the information and this could help users decide whether the source document is worth reading or purchasing. From the knowledge management viewpoint, having a standard rating index tool is probably very useful, but it is very expensive and is not easily available. The concept has been raised but not taken up successfully because of its complexity. Moreover, even if one wants to provide a rating system it is inadvisable to do so because writers would be

biased or subjective in the assessment of their own writing. Furthermore, professional agencies already exist to provide the necessary tools to facilitate the metadata environment, and such tasks require professional training. The question raised was, “Are the authors willing to go that far and even if they are, is it worthwhile”? This led to the general consensus among the professionals that such a utility is more useful in the filtering process by information professionals later than by authors at the source of writing. Thus, at this point, the responsibility is beyond the scope of abstracting.

So the recommendation was for abstract writers to have greater awareness of electronic information systems and policies of abstracting agencies that they are writing for. However, the more technical capabilities like incorporating metadata and other retrieval utilities should be the responsibility of information professionals, thus suggesting that the ESP practitioner’s role should be limited to the abstract writing process.

Desirable qualities that increase visibility and retrievability of abstracts

The FG participants were next asked if there are specific desirable qualities of abstracts that would enhance their chances of being retrieved on information systems. Should abstract writers then slant their writing with the retrieval system in mind?

The first opinion offered was that in an ideal situation it is desirable for the abstract to be produced by a specialist in the discipline, who is also trained in abstract writing. However, this is very difficult to achieve and is also very expensive. Instead, it is more important for writers to understand the reasons why people use abstracts, and that individual users have different needs. This knowledge is especially pertinent when writing abstracts for use on Internet where the audience is much larger, more diverse and less technical.

Based on this awareness and consideration for audience needs, the qualities of readability and exhaustivity are important and desirable. The IS professionals were unanimous in recommending that abstracts should be written in simple and direct

language so that they are easily comprehensible to both experts and non-experts. In other words, writers should concentrate on presenting coherently and communicatively, qualities, which are not yet attainable by machines. In terms of exhaustivity, abstracts should contain sufficient and significant content representation to function effectively as surrogates.

In addition, when writing abstracts the practice of providing significant keywords is recommended as an effective means to ensure better retrievability. Although this practice appears to contradict an earlier stand, it is thought to be relevant here, and the solution offered to ESP teachers is to help create this awareness in their novice-writers, even if they are not able to teach the semantic webs containing specific keywords.

From this perspective, it would appear that the ESP practitioner or language professional has the important dual-role of providing the rhetorical structures and language conventions of abstracts and training novice-writers to produce abstracts of high readability and retrievability. And, where exhaustivity and representation of content are concerned, engaging the collaboration of subject specialists should be the solution if the subject matter is too technical or specialised.

How IS professionals rate linguistic aspects of abstracts

In this section, the FG participants were asked to make recommendations that would synergise with the linguistic concerns of ESP practitioners? Different viewpoints emerged during the discussion on how abstract writers can achieve better results in their writing. To begin this discussion, the term 'linguistic competency' was clarified as referring to lexico-grammatical accuracy of language use. A few participants felt that although language competency is important it is the content representation that is more significant. This is especially true in the case of technical subject matter where the content is valued more than the eloquence of expression. The argument was good language does not necessarily mean good content. However, a note of disagreement was observed especially from one IS professional who is also a journal editor. He considered language fluency and accuracy as equally important if not more important than the content representation in an abstract. This is mainly because the abstract is a tool for communicating brief but significant information to users, and as such the message must be succinctly and effectively expressed if the abstract were

to achieve its primary purpose as surrogate text. ‘Bad language’ may detract readers from wanting to read the abstract, and poor linguistic competence skills have been found to be responsible for abstracts lacking in exhaustive and accurate content. However, it was also agreed that for abstracts to be truly effective the communicative message is more important than mere language or grammatical accuracy. Therefore, effective abstracts must have a good balance of readability and exhaustivity.

The experts then identified several other abstract-writing qualities apart from linguistic competency and content representation that could affect abstracts. The first quality of brevity or word limit could affect the exhaustivity and representation of its source document. The length and content structure could in turn be dependent on the type of abstract and agency requirements. Other constraints could include specific journal guidelines, technicality of subject matter, and agency instructions. However, these are constraints that are not within the control of writers, and thus writers should focus on features that are within their control such as readability and linguistic qualities. From this viewpoint, the participants reiterated that the key is in writing readable abstracts, and to improve readability, the writer would need to write in simple language with clarity, conciseness, and precision, while ensuring good cohesive and coherent text relationships.

When the IS professionals were asked to rank three aspects of the abstract (content presentation, linguistic competency, and exhaustivity or content representation) on a scale of 1-3, with rank 1 being the most important, a consensus was returned in the following order:

- Rank 1 Content: The emphasis is on the significance of content presentation (communicative competence) and content representation (interpretation and selection) of source document. This aspect implies that good content is basically dependent on effective linguistic expression.
- Rank 2 Linguistic competency: This refers to the lexico-grammatical (linguistic competence) accuracy of the abstract. This aspect implies that content representation of source document is less dependent on linguistic accuracy of the text.
- Rank 3 Exhaustivity: This refers to how comprehensive the content of source document is represented, and whether the abstract has captured all its salient information. This aspect is difficult to achieve and measure because of the unique needs of users. This aspect implies that the ESP teacher would have little control over content representation.

Based on the discussion above the following observation is made. While ESP practitioners have made assumptions that linguistic competency is prerequisite for successful abstracts, IS professionals are saying that lack of linguistic competency does not necessarily mean unsuccessful abstracts, but they have nevertheless agreed that linguistic competency enhances the quality of abstracts.

ESP and IS collaboration in abstract writing research and pedagogy

With regards to the feasibility of collaboration, some significant points surfaced. First, the IS professionals felt that it is very important for abstract writers to realise the purpose and type of abstracts to be written, and to appreciate the environment in which the abstract is going to be retrieved. It is also important for writers to abide by the requirements of abstract agencies or journals. For example, in the Web environment writers should realize that they are reaching a much wider and more varied audience, who comprises specialists and laymen. In this environment writers would need to ensure that their abstracts cater to this large readership by writing more explicitly and including more cohesive devices to increase readability. On the other hand, where the audience is highly specialised and the abstract is meant for in-house publications, and where information is currency within the specific discourse community there is less need for too much explicitness. For example there will be no necessity to include non-content words like ‘This paper aims to...’ or ‘The results of this study show...’. Exclusion of the latter cohesive devices or markers saves on words and could enable more information representation of the source document.

Another point raised was when an abstract reaches the IS professional it is in the final product state and if it has been poorly written it would need to be rewritten or modified, which is an expensive exercise. In order to save such cost, the ESP practitioner has the important duty of ensuring that writers produce readable and useful abstracts. A major complaint raised was that abstracts that come to them often do not include relevant or significant content; some even contain information not in the source documents. On this note, IS professionals would also like to see abstracts that highlight the author’s own contributions rather than report on other works, and they believe the ESP practitioner has a role to play here.

From this discussion, it is clear that the roles of ESP practitioners and IS professionals are complementary. At one end, the ESP practitioner encourages and guides the author, who has the advantage of knowing the content better than anybody else, in producing the effective abstract. At the other end, the IS professional evaluates whether the finished products serve their purposes and needs of users when they are put into information systems. But such synergy is possible only if both experts collaborate. Figure 2 demonstrates how the two disciplines can work together.

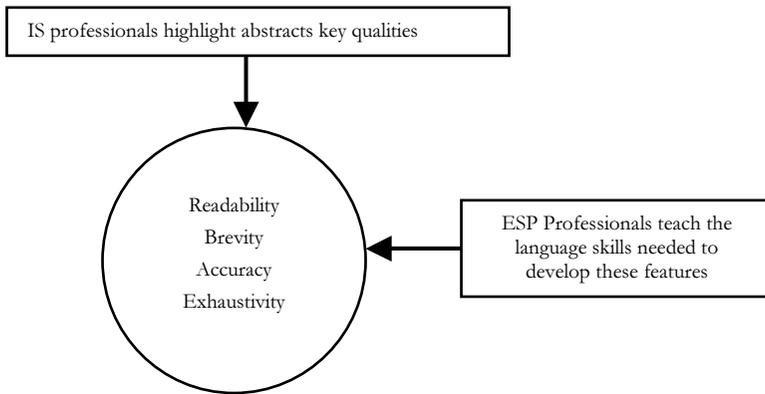


Figure 2. ESP and IS collaboration in abstract research and pedagogy.

For example, the ESP practitioner should develop teaching materials and strategies that help learner-writers to identify relevant content and write concisely and precisely to produce simple, brief, accurate and highly readable surrogates that are acceptable in the information searching environment. Or they should physically collaborate with specialist informants to check on content representation or on technical aspects like accessing semantic webs.

The discussion ended on the note that ESP practitioners should inform novice-writers in areas such as selecting keywords that would boost retrievability, writing abstracts that are friendly to retrieval systems and end-users, and understanding how information systems and information retrieval work. In this collaboration IS professionals would save on time, cost, and the trouble of having to write or rewrite abstracts for authors. The following recommendations were offered for improving the abstract writing pedagogy:

- Abstract writers must realise that good abstract writing means writing communicatively, and making sure that the intended message is clearly and

concisely conveyed. Teach learners to write abstracts that have brevity, simplicity, clarity and accuracy.

- Writers must write abstracts that have high readability and retrievability in the IR system. This is very important, especially when writing for prestigious journals where the abstract is highly depended upon to create the favourable first impression for its source document.
- Abstract writers must be aware of purposes and uses of abstracts, and how different disciplines have different requirements and use different types of abstracts.
- Learner-writers should be encouraged to read each other's abstracts. This practice will help identify information gaps, which writers may miss out. This is a good check on accuracy and exhaustivity of content representation.
- The ESP teacher must first know about the information retrieval (IR) environment, and assist learner-writers in determining suitable descriptors of abstracts so that they can be retrieved by the largest number of readers.

Conclusion

The study adds testament to the perception that ESP teachers are very concerned with learner-writer centered objectives, and spend their energies on coaching learners in the linguistic and structural aspects of writing for specific discourse communities. On the other hand, IS literature shows a different perspective. IS professionals focus on the benefits of using abstracts as surrogate texts in information retrieval, and stress the seriousness of maintaining abstract qualities.

Based on these apparent differences, ESP practitioners would need to expand their scope and direct their learner-writers to the requirements of international standards for abstraction and draw their attention to meeting end-user needs. They should encourage learners to focus more on ensuring content quality, and to view abstracts within the more global context of information seeking and retrieval via information systems, and to aim beyond just producing effective summaries of source articles.

However, we have also seen that IS professionals corroborate with ESP practitioners on valuing readability quality, and stressing the importance of clarity, structure, and brevity of abstracts. In this respect, ESP instructors should retain their role of training learner-writers to write concisely, make precise lexical word choices, use clear

and simple sentences, organise information elements coherently and cohesively, and ultimately produce effective abstracts for information systems.

A major part of the discussion focused on recent impact of technological development in information systems on abstracts and abstract writing. This development has several implications for the teaching of the genre. First, there is the need to heighten learner-writers' awareness in issues related to the use of abstracts in Web and other online environments, and the following issues deserve more consideration and research:

- The increase in the speed of information accessibility or retrieval has significant effects on writing and disseminating information
- Technology in information communication raises the importance of the abstract as an information transfer tool
- Although ESP practitioners have mainly focused on informative abstracts, the shorter and more concise indicative abstract is increasingly preferred on the Web and other online environments.
- The provision of key words derived from semantic groupings has the potential of giving abstracts better visibility to a wider audience, but should it remain the responsibility of the IS professionals and not the ESP practitioners?
- The proposition of using hyperlinks in abstracts could help researchers access specific research data in source documents easily. This potential should be viewed as a useful future development in information retrieval, and it would benefit abstract writers to keep abreast of technological developments.

These issues listed above may or may not necessarily be translated into instructional materials of ESP practitioners because of certain technicalities, but by drawing on the learner-writers' attention to them, will help change their perceptions of abstracts and their potentials. It would add on a new dimension to the learners' perception of the abstract and its role in scholarly communication. It would benefit them to be aware that technology facilities may affect structures and qualities of abstracts in terms of generic structure, length and speed of circulation. Such awareness would prepare learner-writers to write more purposefully, and with this awareness they have two angles to consider: the linguistic and communicative angle as well as the information systems angle. The common linguistic interests shared by both disciplines have useful implications for collaborative work.

REFERENCES

- American National Standards Institution (1979). *American National Standard for Writing Abstracts*. ANSI Z39.14-1979. New York: American National Standards Institute.
- Berkenkotter, C & T. Huckin (1995). *Genre Knowledge in Discipline Communication*, 27-44. Hillsdale, NJ: Lawrence Erlbaum.
- Bhatia, V. K. (1993). *Analysing Genre: Language in Professional Settings*. London & New York: Longman.
- Borko, H. & C. Bernier (1975). *Abstracting Concepts and Methods*. Library and Information Science Series Academic Press.
- Chowdhury G. G. (1999). "Abstracts and abstracting" in *Introduction to Modern Information Retrieval*, 144-157. London: Library Association Publishing.
- Cleveland, D. & A. Cleveland (1983). *Introduction to Indexing and Abstracting*. Littleton, CO: Libraries Unlimited.
- Collison, R. (1971). *Abstracts and Abstracting Services*. Santa Barbara, CA: American Bibliography Center. Clio Press.
- Cremmins, E. T. (1996). *The Art of Abstracting*, 2nd ed. Arlington: Information Resources Press.
- Davis, L. E. (1991). "Student abstract writing as a tool for writing across the curriculum in large introductory-geology courses". *Journal of Geological Education* 39,3: 178-180.
- Dronberger, G. B. & T. Kowitz (1975). "Abstract readability as a factor in information systems". *Journal of the American Society for Information Science* 26,2: 108-111.
- Dudley-Evans, T. (1986). "Genre analysis: an investigation of the introduction and discussion sections of MSc dissertations" in M. Coulthard (Ed.), *Talking about Text*, 128-145. Birmingham: English Language Research.
- Dudley-Evans, T. (1994). "Genre analysis: An approach to text analysis for ESP" in M. Coulthard (Ed.), *Advances in Written Text Analysis*, 219-228. London & New York: Routledge.
- Fidel, R. (1986). "Writing abstracts for free text searching". *Journal of Documentation* 42,1: 11-21.
- Freedman, A. (1994). "Anyone for tennis"? in A. Freedman & P. Medway (Eds.), *Genre and the New Rhetoric*, 43-66. London: Taylor & Francis.
- Gosden, H. (1995). "Success in research article writing and revision: a social constructionist perspective". *English for Specific Purposes* 14,1: 37-57.
- Graetz, N. (1985). "Teaching EFL students to extract structural information about abstracts". In J. M. Ulijn & A. K. Pugh (Eds.), *Reading for Professional Purposes*, 123-135. Leuven: ACCO.
- Hartley, J. M. S. & A. Blurton (1996). "Obtaining information accurately and quickly: Are structured abstracts more efficient?" *Journal of Information Science* 22,5: 349-356.
- Hills, P. J. (1983). "The scholarly communication process". *Annual Review of Information Science and Technology* 18: 99-125.
- Huckin, T. & L. A. Olsen (1983). *English for Science and Technology: A Handbook for Non-Native Speakers*. New York: McGraw-Hill.
- Huckin, T. (2001). "Abstracting from abstracts" in M. Hewings (Ed.), *Academic Writing in Context: Implications and Applications. Papers in honour of Tony Dudley-Evans*, 93-105. Birmingham: The University of Birmingham. University Press.
- ISO 214. (1976). (E) *Documentation-Abstracts for Publications and Documentation*.
- Kan, M. Y., K. R. McKeown & J. L. Klavans (2001). "Domain-specific informative and indicative summarization for information retrieval". *Proceedings of the Document Understanding Workshop* (DUC 2001). New Orleans, USA: September 2001.
- Keogh, T. J. (1994). *The Structure of Abstracts: Stylistic and Structural Elements in 48 Scientific and Technical Abstracts*. PhD Dissertation UMI. AAT 9524462.
- Lancaster, E. W. (1991). *Indexing and Abstracting: Theory and Practice*. London: LA Publication.
- Lorenz, J. G. (1969). "International transfer of information" in C. A. Cuadra (Ed.), *Annual Review of Information Science and Technology*, vol. 4, 379-402 of *Encyclopaedia Britannica*. Chicago: William Benton.
- Marshall, S. (1991). "A genre-based approach to the teaching of report writing". *English for Specific Purposes* 10,3: 3-13.
- Morgan, D. L. (1998). *The Focus Group Guidebook*. Thousand Oaks, CA: Sage Publications.
- Mustafa, Z. (1995). "The effect on genre awareness on linguistic transfer". *English for Specific Purposes* 14,3: 247-256.
- Ngowu, K. N. (1997). "The medical research paper: Structure and functions". *English for Specific Purposes* 16,2: 119-138.
- Paltridge, B. (1997). *Thesis and dissertation writing: Preparing ESL students for research*. *English for Specific Purposes* 16,1: 61-70.
- Pinto, M & F. W. Lancaster (1999). "Abstracts and abstracting in knowledge discovery". *Library Trends* 48,1: 234-248.
- Posteguillo, S. (2002). "El título y el resumen de un artículo de resumen" en I. Fortanet (coord.), *Cómo escribir un artículo de investigación en inglés*. Madrid: Alianza Editorial.
- Posteguillo, S. (1999). "The schematic structure of computer

- science research articles". *English for Specific Purposes* 18,2: 139-160.
- Posteguillo, S. (1996). "A genre-gased approach to the teaching of reading and writing abstracts in computer science" in J. Piqué, J.-V. Andreu-Besó & D. J. Viera (Eds.), *English in Specific Settings*, 47-57. Valencia: NAU Llibres.
- Rowley, J. (1982). *Abstracting and Indexing*. London: Clive Bingley.
- Salbiah S. (2000). "The genre of abstracts of conference proceedings". Paper presented at LSP Seminar, November 2000. Johor Bharu, Malaysia.
- Salager-Meyer, F. (1990). "Discourse flaws in Medical English abstracts: A genre analysis per research and text type". *Text* 10,4: 365-384.
- Santos, M. B. D. (1996). "The textual organization of research paper abstracts in applied linguistics". *Text* 16,4: 481-499.
- Sionis, C. (1995). "Communication strategies in the writing of scientific research articles by non-native users of English". *English for Specific Purposes* 14,2: 99-113.
- Swales, J. M. (1981). "Aspects of article introductions." *University of Aston ESP Research Reports*, No. 1. Birmingham: Language Studies.
- Swales, J. M. (1984). "Research into the structures of introductions to journal articles and its applications to the teaching of academic writing". In R. Williams, J. Swales & J. Kirkman (Eds.), *Common Ground: Shared Interests in ESP and Communication Studies*, 77-86. Oxford: Pergamon Press, ELT Documents 117.
- Swales, J. M. (1985). *Episodes in ESP*. Oxford & New York: Pergamon Institute of English.
- Swales, J. M. (1986). "A genre-based approach to language across the curriculum". In M. L. Tickoo (ed), *Language Across the Curriculum. Selected papers from the RELC Seminar on 'Language Across the Curriculum'*. Singapore, 22-26 April 1985, 10-22. Singapore: SEAMEO Regional Language Centre.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Tenopir, C. (1985). "Full text database retrieval performance". *Online Review* 9,2: 149-164.
- Tenopir, C & P. Jacso (1993). "Quality of abstracts". *Online* 17,3: 44-55.
- Wheatley, A. & J. C. Armstrong (1997). "Metadata, recall, and abstracts: can abstracts ever be reliable indicators of document value"? *Aslib Proceedings* 49,8: 206-213.

Soon Keng Chan is a lecturer at Nanyang Technological University (NTU), Singapore. She teaches English proficiency, professional writing, interpersonal communication, and negotiating skills to undergraduates. She conducts writing workshops for foreign postgraduates from China and Indonesia. She is working towards a dissertation on ESL research scholars' abstract writing profiles.

Schubert Foo is Professor at the School of Communication and Information at Nanyang Technological University (NTU). He has taught courses in the Divisions of Computer Technology, Software Systems and Information Studies. He has published in his research areas of multimedia technology, Internet technology, multilingual information retrieval, digital libraries and knowledge management.

Appendix

Focus Group Questions

Current technological advancements in information dissemination and retrieval are posing an additional challenge to English for Specific Purpose (ESP) practitioners who teach academic writing. Is the current ESP focus on teaching rhetorical structures and linguistic conventions of genres enough? Have ESP practitioners a role in helping novice-writers meet the needs of abstracts for the web environment?

1. Studies have shown that electronic abstracts (e.g. those in Internet directories and Internet gateways) are different from traditional online database abstracts derived from print-source environment (LISA, ERIC). Based on your experience as an information provider and manager, what is your opinion? How far do you think this claim is true?
2. In this environment of technology, how can **author abstracts** meet the requirements of electronic IR systems so that they would stand the best chances of being retrieved?
3. Specifically, are there certain vocabulary terms or data about the abstract that would make it more responsive to IR systems?
4. Should or can standard information organization and retrieval tools of IR, e.g. subject headings, thesaurus, and controlled vocabulary, be **taught** so abstract writers will better cater to the requirements of IR systems?
5. Some researchers have said that, an ideal 'Internet abstract' should include metadata such as user guidance, assessment of authority, discussion of physical attributes, judgements of quality, or pointers to alternative sources. (Wheatley, 1997). To what extent would you agree?
6. How significant is **linguistic accuracy** of an abstract to information experts as selection criteria for retrieval systems? Please explain.
7. How does linguistic accuracy **compare** with **content representation and exhaustivity** of the abstract in the selection process? Please explain.
8. In your opinion, what common concerns do information experts share with language providers about the art of abstract writing?
9. How can language providers collaborate with information experts so that they can be better informed to effectively cater to the changes in technology and information systems?