Science, Systemic Functional Linguistics and Language Change

Science, Systemic Functional Linguistics and Language Change:

A Festschrift for David Banks

Edited by

Shirley Carter-Thomas
and Clive E. Hamilton

Science, Systemic Functional Linguistics and Language Change:
A Festschrift for David Banks

Edited by Shirley Carter-Thomas, Clive E. Hamilton

TABLE OF CONTENTS

LIST OF CONTRIBUTORS

Shirley Carter-Thomas: Institut Mines-Télécom Business School

Carmela Chateau Smith: Université de Bourgogne

Dacia Dressen-Hammouda: Université Clermont Auvergne

Pascaline Faure: Sorbonne Université

Lise Fontaine: University of Cardiff

Jacques François: Université de Caen Normandie

Clive Hamilton: Université de Paris

Janet Ormrod: IMT Atlantique (ex Télécom Bretagne)

Anca Christine Pascu: Université de Bretagne Occidentale

Ghislaine Rolland-Lozachmeur: Université de Bretagne Occidentale

Elizabeth Rowley-Jolivet: Université d'Orléans

Mohamed Saki: Université de Bretagne Occidentale

Gordon Tucker: University of Cardiff

INTRODUCTION

SHIRLEY CARTER-THOMAS
AND CLIVE HAMILTON

1. David Banks: Introductory remarks

David Banks has been a leading figure in linguistics, both within France and internationally, for a number of years. He is the author of a prolific body of research which has influenced several generations of researchers, and has also been successful in bringing together linguists of different theoretical backgrounds, languages and research domains. Emeritus Professor of English linguistics at *Université de Bretagne Occidentale* in Brest, where he has been based for over three decades, David Banks is originally from Newcastle-upon-Tyne. He has however been living outside the UK since 1975, initially in Iraq, and subsequently in France. He holds a degree in philosophy from the University of Cambridge, a doctorate in English linguistics from the *Université de Nantes*, and an HDR (*Habilitation à Diriger des Recherches*) from the *Université de Bordeaux 2*.

David Banks founded the French systemic functional linguistics group AFLSF (*Association Française de la Linguistique Systémique Fonctionnelle*), and was its Chairman for fourteen years. He was also formerly Head of the English department and Director of the ERLA (*Equipe de Recherche en Linguistique Appliquée*), the applied linguistics team at *Université de Bretagne Occidentale,* and organised regular conferences at his university, most of which gave rise to collective publications on a vast array of topics. The different papers assembled in this book by fellow scholars and colleagues of David Banks are intended as a tribute to these and other achievements, in recognition of the seminal role David Banks has played in linguistics. It is also hoped that other younger colleagues and scholars less familiar with his work will take an interest in the issues raised and be encouraged to delve further.

The title of the present book, "Science, SFL and Language Change: A Festschrift in honour of David Banks", refers to three of the main themes that have run through his work over the years: scientific language, Systemic

Functional Linguistics (SFL), and diachrony and language change. Although these three strands have often been amalgamated in David Banks's work, we propose to briefly describe his interest in the three separately, before going on to outline the organisation of the book.

2. Science, SFL and Language Change

One of David Banks's main passions has been for the language of science, culminating in the publication of *The Development of Scientific Writing: Linguistic Features and Historical Context* (Banks 2008), for which he received the European Society for the Study of English Language and Linguistics Book Award in 2010. This work, after first discussing previous studies of the development of scientific language, including Halliday's exploration of the topic (Halliday 1988), traces the development of scientific writing as a genre, in terms of its linguistic features, from Chaucer's "Treatise on the Astrolabe" to a corpus of texts taken from the Philosophical Transactions of the Royal Society, covering the period 1700 to 1980. Basing his analysis on a close study of a range of linguistic features (first person pronouns, nominalisation, grammatical metaphor, passive forms and thematic structure), his conclusions point to some different patterns of development in the physical and biological sciences, such as the increased influence of mathematical modelling on texts in the physical sciences from the late 19th century onwards.

David Banks's focus on the grammatical features of scientific writing is a prominent trait of much of his work. In his earlier work, this emphasis was also tightly bound up with the pedagogical concern of providing a detailed description of scientific journal articles for the training of non-Anglophone scientists. Using a small set of articles in oceanography (1991, 1993, 1995 *inter alia*), David Banks surveyed a variety of grammatical features such as the use of personal pronouns, agentless passives, hedging patterns, etc. Rather than considering scientific texts from a rhetorical and sociolinguistic viewpoint, his starting point here was the text's linguistic specificities, those present as well as those absent (cf. Banks 1995).

David Banks's interest in applied linguistics in France can also be seen in his consistent support for the GERAS[1] conferences over the years, his membership of the editorial board of the *ASp* journal and his editorship (in the pre-Internet days) of the ESP France newsletter. He has done a great deal to promote LSP and ESP studies in the French linguistics landscape,

---

[1] *Groupe d'Etude et de Recherche en Anglais de Spécialité* (English for Specific Purposes Study and Research Group).

not only though his own work but also though his PhD supervision, participation in committees and the support he has always so generously provided to his colleagues.

The promotion of Systemic Functional Linguistics in France is also among one of David Banks's most significant achievements. As well as teaching SFL in his courses, disseminating SFL through his journal articles (cf. for example Banks 2002) and his recent book (2019), he is the author of an abridged SFL grammar of English published in French (2005), *Introduction à la linguistique systémique fonctionnelle de l'anglais.* With a foreword by Michael Halliday, this very readable grammar covers several aspects of SFL theory such as transitivity analyses, thematic structure and the interpersonal metafunction, and has been instrumental in enabling French linguists (students and teachers alike) to gain a better understanding of SFL theory. David Banks has in addition co-edited books bringing together SFL analyses of the French language on various features (2009, 2016). In an effort to build bridges between theories, he has also drawn parallels between SFL and the French theory of enunciative operations, based on the work of Antoine Culioli (Banks 2004).

The study of language as it develops over time is perhaps the linguistic approach that has come to best represent the work of David Banks, particularly when combined with the study of scientific discourse. In recent years David Banks has pioneered a diachronic approach to the study of English (and other languages) for specific purposes, or "diachronic ESP" (Banks 2005b; 2012; 2016). He argues that *"understanding the way specialised texts have developed can help us understand why they are the way they are today"* (Banks 2012, 55). Scientific language necessarily changes in function of the context of scientific research in which it is produced, in order to meet the communicative needs of its practitioners.

In another recent book on the birth of the scientific article (Banks 2017), he uses the method to compare and contrast the development of research articles taken from the first two academic periodicals in Britain and France, from their creation in 1665 until the end of the seventeenth century. The focus on the historical context of the late seventeenth century in the two countries enables him to place the texts solidly in the context from which they derived. It is firstly shown how this context influenced the different editorial decisions made by the respective editors of the *Journal des Sçavans* and the *Philosophical Transactions*, which led in turn to the texts they published also displaying differing linguistic features. David Banks makes a convincing case for the contrastive analysis of various linguistic features studied in the two journals, such as type of theme, nominalisations,

expressions of modality or process type, being directly explicable in terms of the historical context (Banks 2012).

The emphasis on three main themes in David Banks work should not be taken as an indication that his interests have only been limited to these topics. In addition to scientific texts, he has also analysed the language of the Bible, fiction, and the press (cf. Banks 1999). His work embraces lexico-grammar and the study of text, both from synchronic and diachronic viewpoints. He has also edited volumes on a variety of subjects: coordination and subordination, tense and aspect, poetic language, political texts and propaganda (cf. Banks 2005c; 2008).

3. Presentation of Contributions

The contributions in this book were chosen to provide a picture of the main fields of interest of David Banks, as outlined in section 2: SFL, the language of science and language change. The perspective offered is thus quite unique in the coverage it proposes of these three areas, which are rarely treated together. The eleven chapters can all be read separately, but the numerous points of convergence and inter-connections between the chapters and sections contribute to the book's overall coherence.

The authors are experienced linguists and/or LSP practitioners, all with considerable experience of teaching and research in higher education in France and the UK. Some are colleagues of David Banks from *Université de Bretagne Occidentale* in Brest, others are colleagues from the various scholarly associations David Banks have been associated with, or who have met him regularly in research contexts over the years. Several of the contributors have worked in the French university system and the book could therefore be of particular interest for scholars working in France or other European contexts, or for those with interests in contrastive linguistics. Four of the eleven chapters are in French.

The papers proposed in the first section all consider aspects of ESP in scientific discourse. **Carmela Chateau Smith** (Université de Bourgogne) focuses on the predominance of English as the language of scientific communication. After exploring the history of English in this role, she discusses some of the problematic issues linked to the notions of "native" and "non-nativeness" in academic writing in the sciences. The article ends with some suggestions for how language assistance can be best optimised by ESP teachers. Taking also scientific discourse as her starting point, **Elizabeth Rowley-Jolivet** (Université d'Orléans) focuses on the question of lexical density. She shows that although lexical density and the associated informational load are generally considered to be very high in scientific

discourse (cf. Halliday 1985), it is also important to take into account other factors, such as genre and mode. The analysis of a corpus comparing seven scientific research genres - two spoken (three-minute thesis presentations, conference talks), and five written (research articles, thesis and article abstracts, scholarly journal editorials, laboratory protocols, conference slides) - supplemented by analysis of their verbal density and subject pronoun use, produces results that show great variation. Whereas for the two speech genres lexical density stands at around 48%, it rises to 75% in writing, and also displays frequent correlations with the other two features examined. **Pascaline Faure** (Sorbonne Université) focuses in her chapter on medical English. She shows how English medical terminology is motivated on the one hand by the nature and constraints of Western medical practice (brevity and medical secrecy) and on the other hand by Anglo-Saxon culture (merit, science, travel, humour, food, and lifestyle). After providing examples to illustrate these two main influences, she goes on to consider the case of recently created medical terminology linked to the emerging field of 4P (or preventive) medicine, and to the possible influence of new types of medical cultures associated with this movement.

The second section of the book focuses on diachronic approaches in the analysis of specialised discourse. In her chapter, **Dacia Dressen-Hammouda** (Université Clermont Auvergne) explains the importance of diachronic analysis as a methodology for identifying and exploring the implicit content of disciplinary writing. Drawing on previous work carried out in the discipline of geology (Dressen-Hammouda, 2014), she shows that by situating current text features within past contexts, it is possible to identify implicit or occluded features of contexts that have become less explicit over time. The author claims that this socio-historical insight into the implicit features of writing has a considerable contribution to make in the field of teaching methods in specialised writing. **Janet Ormrod** (ex-Telecom Bretagne) focuses on the popularisation of science in 19th century France. Basing her analysis on a text intended for a lay audience (from the "Bibliothèque utile") with an article on same topic that appeared in the scientific journal, *Journal des Sçavans*, her study identifies some of the characteristics that distinguish the two genres: "low science" and "high science. **Mohamed Saki** (Université de Bretagne Occidentale) in his chapter examines the interpersonal and evaluative dimensions of metadiscourse in six auctorial prefaces to English novels written in the 18th, 19th and 20th centuries. The analysis seeks to shed light on the permanent and changing characteristics of the uses of metadiscourse in the preface subgenre. **Ghislaine Rolland-Lozachmeur (**Université de Bretagne Occidentale) turns her attention to the medical domain (cf. chapter by Pascaline Faure

above). Her contribution analyses the definitions of the word "cancer" in language dictionaries and medical works, from the 18th century up to today's health advice and information websites. The analysis suggests a great deal of consistency in the way the term is used, with the disease being frequently associated with a lexicon of combat.

The chapters in the third section, as well as focusing on language change, also raise issues related to linguistic theory. SFL approaches to language and language change form the basis of the chapters by **Lise Fontaine** (University of Cardiff) and **Gordon Tucker** (University of Cardiff). Starting from the premise that seeds of potentially permanent language change can be observed over the course of just a few years, Gordon Tucker focuses on a case of possible contemporary language change. The paper charts evidence, in corpus data, for the recent increase in the use of the preposition "into", instead of "in", as in the expression "We shall shortly be arriving into Reading" and offers possible explanations of why this apparent change is taking place. Lise Fontaine in her contribution analyses the linguistic complexity of the very topical neologism, *Brexit*. As a term, Brexit represents a complex nominal since it exploits nominal grammatical resources whilst expressing process semantics. She argues that while the newness of *Brexit* prevents the concept of (in)congruence from being operational, the SFL concept of construal in the other hand offers a potentially useful approach for empirical evidence of its semantics, with its construal being that of a hypothetical event with negative semantic prosody. **Anca Christine Pascu** (Université de Bretagne Occidentale) analyses the relationship between natural, iconic and symbolic languages. After tracing the development of linguistics and logic as disciplines, she shows how the distinction between them can also be used as a parallel to distinguish iconic language and symbolic language. Drawing on her background in computer science and natural language processing, her aim is to establish some characteristics of these three types of language and their combined impact on the construction of knowledge. In the final chapter, **Jacques François**, (Université de Caen Normandie) studies the evolution of 19th century German linguistic discourse, based on a close analysis of three linguists. In the first half of the century, illustrated by the writings of Wilhelm von Humboldt and Franz Bopp, the linguistic discourse is directed at peers. However, by the end of the century, with the rapid expansion of universities in Germany following the country's unification in 1870, this discourse takes on a more applied educational dimension, illustrated in the excerpt from Meyer-Lübke. By distinguishing between different types of relationships between clauses, the author shows how SFL is a useful tool for comparatively assessing how the syntactic structure of these discourses evolved.

∗∗
∗

It remains for us to hope that readers will enjoy reading the contributions, which reflect the great respect in which David Banks is held, as much as we have enjoyed preparing this volume as a tribute to his achievements.

References

Banks, David. 1991. "Some observations concerning transitivity and modality in scientific writing." *Language Sciences*. 13(1). 59–78.

Banks, David. 1993. *Writ in Water. Aspects of the Scientific Journal Article*. Brest: E.R.L.A. Université de Bretagne Occidentale.

Banks, David. 1995. "There is a cleft in your sentence: Less common clause structures in scientific writing." Asp. 7-10: 3-11. https://journals.openedition.org/asp/370

Banks, David. 1999. "Decoding the information structure of journalistic clefts" *Interface: Journal of Applied Linguistics,* 14.1: 3-24.

Banks, David. 2002. "Systemic Functional Linguistics as a model for text analysis." *ASp*, 35-36: 23-34. https://journals.openedition.org/asp/1584

Banks, David. 2004. "Anglophone systemicists and French enunciativists: Shall the twain never meet?" *Language Sciences* 26(4): 391-410

Banks, David. 2005a, *Introduction à la linguistique systémique fonctionnelle de l'anglais.* Paris: L'Harmattan*.*

Banks, David. 2005b. "On the historical origins of nominalized process in scientific text." *English for Specific Purposes* 24/3, 347–367.

Banks, David (Coord.) 2005c. *Aspects linguistiques du texte de propagande*, Paris, L'Harmattan.

Banks, David. 2008*. The Development of Scientific Writing: Linguistic Features and Historical Context*. London: Equinox.

Bank, David (Coord.) 2008. *La Langue, la linguistique et le texte religieux*. Paris, L'Harmattan.

Banks, David, Simon Eason, Janet Ormrod (Eds). 2009. *La linguistique systémique fonctionnelle et la langue française.* Paris, L'Harmattan.

Banks, David. 2012. "Diachronic ESP: at the interface of linguistics and cultural studies", *ASp*, 61: 55-70.

Banks, David. 2016. "Diachronic aspects of ESP." *ASp*, 69: 97-112. https://journals.openedition.org/asp/4812

Banks, David, Ormrod, Janet (Coord.) 2016. *Nouvelles études sur la transitivité en français*. France. L'Harmattan.

Banks, David. 2017. *The Birth of the Academic Article: Le Journal des Sçavans and the Philosophical Transactions*, 1665-1700. Sheffield: Equinox

Banks, David. 2019. *A Systemic Functional Grammar of English. A simple introduction.* New York & London, Routledge.

Dressen-Hammouda, Dacia. 2014. "Measuring the Voice of Disciplinarity in Scientific Writing: A Longitudinal Exploration of Experienced Writers in Geology*." English for Specific Purposes* 34: 14–25.

Halliday, Michael A.K.1985. *Spoken and written language.* Oxford: Oxford University Press.

Halliday, M.A.K. 1988. "On the language of physical science". In M. Ghadessy. (Ed.), *Registers of Written English: Situational factors and linguistic feature*s. Pinter: London, 162–178. [reprinted in Halliday, M.A.K. & J.R. Martin 1993. *Writing Science. Literacy and discursive power.* The Falmer Press: London, 54–68].

PART I:

LANGUAGES FOR SPECIFIC PURPOSES

CHAPTER ONE

THE LANGUAGE OF SCIENCE
AND THE NATIVE SPEAKER:
NATURE, NURTURE, AND NORMS

CARMELA CHATEAU SMITH

Introduction

Two frequent themes in the work of David Banks are the importance of diachronic studies in English for Specific Purposes (ESP), and the difference in worldview between English and French. This study combines these themes by exploring the history of English as an international language of scientific communication, together with the notion of "native speaker".

In the early twenty-first century, English is acknowledged as the international language of science, yet this predominance is comparatively recent. The first section of this study, the Nature of scientific language, traces the history of languages used for international communication. The second section, Nurture, explores a pilot corpus of 27 articles providing advice on effective writing for publication. The final section, Norms, examines aspects of linguistic monopoly in academia today, through a case study based on reviewers' remarks. Optimising linguistic assistance for all, from first-year students to experienced reviewers, remains a key goal for ESP teachers within the scientific community.

1. Nature: the language of science over time

The history of English as an international language for scientific communication is linked to the history of science. Isaac Newton (1643–1727) was a "natural philosopher", as underlined by David Banks in his comparison of Newton's *Opticks* and Christiaan Huygens' *Traité de la Lumière*, an exploration of "how scientific English came into being in the

late seventeenth century" (Banks, 2004: 2). The term "scientist" was coined much later, in 1834, by William Whewell (Ross, 1962: 71).

This section is divided into four parts, covering the impact of Greek and Latin until the late seventeenth century, then French in eighteenth- and nineteenth-century Europe, followed by Esperanto and the quest for a neutral international language in the early twentieth century, and finally the place of English in twenty-first century academia.

1.1 The influence of Greek and Latin

Many languages have spread beyond national borders, but Greek is "the European language with the longest recorded history" (Horrocks, 2010: 10). Much modern scientific vocabulary is based on Greek roots, by direct imports, and by imports from Greek via Latin, so that "over 60 percent of all English words have Greek or Latin roots; in the vocabulary of the sciences and technology, the figure rises to over 90 percent." (Green, 2015: 12). Similar meanings are expressed by words of different origin, e.g. Greek *gé* and *logos* are found in "geological"; Latin *terra* produces "terrestrial"; Old English *eorþe* gives "earthly". For other words (e.g. *agony*), "The route of transmission is from Greek to Latin to French." (Stockwell & Minkova, 2001: 51).

In Britain, the use of Latin began during Roman occupation (AD 43 to AD 410), and continued long after the fall of the Roman Empire, notably as the language of the Catholic Church. Elsewhere in Europe, Latin was chosen in 1479 as the language for the newly founded University of Copenhagen, since its purpose was "essentially religious" (Mortensen & Haberland, 2012: 177). Although John Wycliffe's English Bible was hand-printed as early as 1382, the Church of England's fundamental English texts first appeared during the Tudor Reformation (*Book of Common Prayer*, 1569; *King James Bible*, 1611)

The Scientific Revolution, during the Stuart period (1603–1714), saw the discoveries of Kepler, Galileo, Harvey, Hooke, Boyle, and Newton, together with the works of Bacon, Descartes, Hobbes, and Locke, amongst others. In Lyell's *Principles of Geology*, a footnote about Lazarro Moro's *Sui Crostacei ed altri Corpi marini che si trovano sui Monti* (1740) underlines the difficulty of international scientific communication at that time:

> Moro does not cite the works of Hooke and Ray, and although so many of his views were in accordance with theirs, he was probably ignorant of their writings, for they had not been translated. As he always refers to the Latin edition of Burnet, and a French translation of Woodward, we may presume

that he did not read English. (Lyell, 1830: 42).

A pilot study by David Banks (2005: 355) concluded that "Latin cannot be ignored as a contributing factor in the emerging scientific language towards the end of the 17th century." Some new words combined Greek and Latin roots infelicitously, as Lyell (1833: 53) pointed out:

> I have been much indebted to my friend, the Rev. W. Whewell, for assisting me in inventing and anglicising these terms, and I sincerely wish that the numerous foreign diphthongs, barbarous terminations, and Latin plurals, which have been so plentifully introduced of late years into our scientific language, had been avoided as successfully as they are by French naturalists, and as they were by the earlier English writers, when our language was more flexible than it is now. But while I commend the French for accommodating foreign terms to the structure of their own language, I must confess that no naturalists have been more unscholar-like in their mode of fabricating Greek derivatives and compounds, many of the latter being a bastard offspring of Greek and Latin.

1.2 The role of French, from 1066 onwards

After the Norman Conquest in 1066, Norman French became the language of the aristocracy, while Latin remained the language of religion, leaving English as the language of the common people. England and France coexisted as rival powers for many centuries, with much greater linguistic transfer from French to English. Several authors[1] have shown that the richness of English vocabulary stems from these three strands, with the persistence of quasi-synonyms derived from Latin, French, and English roots, e.g. secure, firm, and fast (Jackson & Amvela, 2007: 41).

In addition to Greek and Latin, academics were also expected to know vernacular languages; many scholars travelled extensively, as the following quotation shows:

> During the 17th century, the German universities were of importance, along with the Dutch ones. Universities in England and France, and even in Italy and Spain, were also visited. The tradition of peregrination was upheld in the eighteenth century as well, though to a lesser degree. It is quite possible, though disputed, that the famous Swedish scholar Linnaeus, for example, bears witness to the learning of languages abroad; that his travelling pupils were versed in languages is well documented. (Gren-Eklund, 2011: 57)

---

[1] For a general overview, see Minkova & Stockwell, 2006.

The transition from Latin to the vernacular in scientific writing occurred progressively during the seventeenth century. Galileo wrote *Sidereus Nuncius* in Latin in 1610; his later works were written in Italian. Descartes wrote the French axiom "Je pense, donc je suis" in 1637, while his more famous "Cogito, ergo sum" appeared later, in 1644. Henry Oldenburg's *Philosophical Transactions* (1665) often published texts translated from Latin or French, perhaps because the journal was "open to everyone" (Banks, 2009: 3). Newton wrote *Philosophiae Naturalis Principia Mathematica* (1687) in Latin, while the later *Opticks* (1704) was written in English.[2] During the eighteenth century, "Saussure's native French was by far the most important [language] in the world of the sciences." (Rudwick, 2005: 29).[3]

Language learning may have been adversely affected by the French Revolution and the Napoleonic Wars that followed, preventing an entire generation of young men from undertaking their Grand Tour, previously considered an essential part of a gentleman's education (Black, 2003). In 1795, Hutton included many French quotations in his *Theory of the Earth*, which was often judged "unreadable" by his peers (Leddra, 2010: 115). By 1830, Lyell would provide the general reader with translations and paraphrases. Dufief, the author of a French language manual, stated that "the English and Americans, when they begin the study of the French language, may be considered as *tabulae rasae*." (1831: iv).

Long before the French Revolution (1789), D'Alembert pessimistically predicted:

> Thus, before the end of the eighteenth century, a philosopher who would like to educate himself thoroughly concerning the discoveries of his predecessors will be required to burden his memory with seven or eight different languages. And after having consumed the most precious time of his life in learning them, he will die before beginning to educate himself. The use of the Latin language, which we have shown to be ridiculous in matters of taste, is of the greatest service in works of philosophy, whose merit is entirely determined by clarity and precision, and which urgently require a universal and conventional language. It is therefore to be hoped that this usage will be re-established, yet we have no grounds to hope for it. (1751, translated by Schwab, 1963: 92–93)

D'Alembert's linguistic hopes were to be renewed after "the war to end all wars" (World War I, 1914–1918), with the quest for a new, more neutral language of peace.

---

[2] *Optice*, the Latin translation by Clarke, was published in 1706.
[3] Horace Bénédict de Saussure, the Swiss geologist (1740–1799).

1.3 The case for Esperanto

Ludwik Lejzer Zamenhof invented Internacia lingvo (International language) in 1887 "to unite the world through a common language" (Patterson & Huff, 1999: 444). Doktoro Esperanto (Dr. Hopeful) was in fact Zamenhof's pseudonym (Gledhill, 2000: 4). In 1933, E. L. Thorndike estimated that Esperanto could be learnt in one-fifth of the time it took to learn French, German, Italian, or Spanish (Fettes, 1997: 152). Eugen Wüster, the terminologist, suggested that it would take six times more energy to learn English than to learn Esperanto (Wüster, 1934). Despite this ease of acquisition, conservative estimates indicate only 40,000 fluent speakers of Esperanto (Gledhill, 2000: 10).

Comparative figures for English, French and Esperanto are presented in Table 1, based on data collected by Sydney S. Culbert and published in the *World Almanac and Book of Facts* in 1994 (Comrie, 1996: 36–37). Lindstedt (2010: 69) estimated that "there are perhaps one thousand first-language speakers".

Table 1: Number of 1st or 2nd language speakers

Language	1st language	2nd language	Total
English	322,000,000	148,000,000	470,000,000
French	72,000,000	54,000,000	124,000,000
Esperanto	1,000	2,000,000	2,001,000

In contrast with English and French, other major languages "are more restricted geographically, e. g. Spanish primarily to Spain and Latin America, Arabic to the Middle East and North Africa, Russian to the former USSR" (Comrie 1996: 37). The Universal Association of Esperanto has members in 120 countries (UEA, 2016). Google Maps and Chatterplot show Esperanto speakers in all time zones. Esperanto could therefore legitimately rank as a geographically widespread language, but with limited speaker density.

France is, surprisingly, the country spanning the greatest number of time zones (Table 2). "This unusual span is due to France's scattered national territories. The areas in French Polynesia in the Pacific Ocean are mainly responsible for this." (*India Today*, October 28, 2015). If the analysis were to be based on the number of time zones outside the mainland area, the United Kingdom would rank second.

Table 2: Number of Time Zones (Source: *India Today*, 2015)

Country	France	USA	Russia	UK	Australia
Mainland	1	6	11	1	3
Other	11	5	0	8	5
Total	12	11	11	9	8

In 1922, the League of Nations[4] proposed to facilitate the adoption of Esperanto as their working language by teaching it in schools, requiring ten votes in favour of the motion. Delegate Gabriel Hanotaux voted against the proposal, probably fearing that French would be endangered by the spread of Esperanto, and this single negative vote was sufficient for the motion to be rejected (Tafalla Plana, 2010: 139). A key opportunity to promote Esperanto as a major international language was lost, yet the international influence of French declined, as English gradually became the most widely used international language.

1.4 The ubiquity of English

Among the reasons for the spread of English are its capacity for integrating words from other languages, and the fact that there is no one variety of English with higher status than the others. "English is thus a language whose native speakers, especially those involved in international communication, are in general tolerant of a wide range of variation and recognize the existence of a number of non-native speaker varieties." (Comrie, 1996: 49). Monolingualism may also explain this tolerance: in 2016, 65.4% of working-age adults in the UK reported that they knew no foreign languages, while the figure for France was only 39.9% (Eurostat, 2018).

One drawback of monolingualism is that it may limit access to scientific knowledge, and even distort knowledge of the history of science. Earth Science was often interpreted through the prism of locally accessible data, as the following reference to "British geology" indicates.

> It might seem (to English-language readers) that Wernerism was decisively defeated by Hutton and British geology in general. However, while the rise of organized field-mapping, described in Chapter 5, was a notable expression of 'English-language geology', the greater part of the basic work in petrography was carried out on the Continent, particularly in France and Germany. (Oldroyd, 1996: 196).

---

[4] http://vortareto.free.fr/argumentaire/sdn/sdn.pdf (Accessed 3 February, 2019).

Translating scientific texts could provide a solution to the problem of monolingualism, but translation and stimulus meanings may not always coincide (Quine, 1960: 40). Wegener's "*Theorie der Kontinental-verschiebung*" was initially translated as the "theory of continental displacement", but it is much more widely known under the term "continental drift", despite or perhaps even because of its negative semantic prosody (Chateau, 2010). Problems may also arise from the general tendency towards lexical simplification in translated texts (Xiao, 2010: 29). In the absence of accurate, culturally relevant translations, the entire paradigm of normal science may become skewed (Kuhn, 1970).

An analysis of the languages cited as references in 139 sedimentological papers published by Elsevier in 1964 led to the following conclusions:

(1) Russian and English were the two major scientific languages, but with little contact between the two areas, except for "a small proportion of translated literature".
(2) French (or German) was seldom read outside France (or Germany).
(3) British sedimentologists read more languages (especially German) than their North American colleagues, who tended to rely on "a smaller number of scientific periodicals".
(4) In North America, older scientists read more languages than the younger generation.
(5) Older Dutch sedimentologists read more French than the younger generation, while younger German sedimentologists read more languages than the older generation. (adapted from Manten & van Dilst, 1965: 22–23)

In 1967, a similar survey by the British National Lending Library concluded that:

The foreign-language barrier facing British scientists is a function of a) the amount and value of foreign scientific literature produced; b) the linguistic ability of the people wishing to read it; and c) the availability and effectiveness of translation services." (Wood, 1967).

Barely forty years later, the barrier to understanding "foreign scientific literature" has almost disappeared, but not through greater knowledge of other languages. In the twenty-first century, the majority of the scientific literature in many fields tends to be published directly in English, which has obliged many professional translators to adapt to this situation, often by providing editing rather than translation services.

Although English, French, and German were used with similar frequency to write scientific articles at the end of the nineteenth century, French has been in constant decline since that time, according to data

collated by Ammon (2012: 338). The use of German increased quite rapidly until the aftermath of the First World War, even overtaking English, but has declined steadily since 1920. The peak in the use of Russian (from 1950 to 1988) is probably linked to geopolitical events, such as the Cold War, the founding of Solidarność, and Gorbachev's dual programme of Perestroika and Glasnost.

Many academics deplore this decline in multilingualism, yet it is undoubtedly true that the effort to master one foreign language is far less than that required for the mastery of the seven or eight languages that D'Alembert thought would be necessary should all scientists use the vernacular for their publications. As Marcia McNutt (2013: 13), Editor-in-Chief of *Science* pointed out: "Even the most brilliant scientific discovery, if not communicated widely and accurately, is of little value." Although it has often been an uphill struggle, at times of paradigm shift, for the authors of controversial theories to gain recognition by the scientific community, as was the case for Wegener and continental drift (Chateau, 2014), it is undoubtedly necessary today for scientists to communicate their discoveries, and a recent report by the American Sociological Association has gone so far as to suggest that making scientific discoveries known to the general public is even more important than communication within the scientific community (ASA, 2016).

2. Nurture: Advice to authors

In order to communicate both widely and accurately, some degree of mastery is necessary. In the field of scientific communication, the peer-review system is seen as a guarantee of quality. In the field of communication in English, the educated native speaker is considered to provide a model worthy of imitation. This can sometimes lead to problems of unwitting plagiarism.

This section explores the notion of native speaker, and then introduces a pilot corpus of Advice to Authors, to draw some conclusions about the constraints that apply to the concept of English as an international language for written scientific communication, which is not generally shown the tolerance accorded to oral communication (Ammon, 2012: 350).

2.1 Defining the model

In order to imitate the native speaker model, it is necessary to define the notion of native speaker. Swales (1985) proposed a series of criteria to differentiate between "Native Speaker (NS)" and "Non-Native Speaker

(NNS)" status, with plus or minus scores of up to 12 points (which would perhaps be judged less than politically correct by today's standards). The terms proposed in this section are: "NES" (Native English Speaker), NFS (Native French Speaker), and NNES (Non-Native English Speaker).

1. R (1 point) Last name (including Anglicization etc.),
2. RRR (3 points) Institutional affiliation (assumed to be permanent unless otherwise indicated),
3. R (1 point) All citations to English language publications,
4. RR (2 points) First name of author (John, Jean, Juan, Giovanni, Johann, etc.; surprisingly useful evidence when available, especially useful for identifying the language status of married women),
5. RR (2 points) All self citations to English language publications,
6. RRR (3 points) NS/NSS Status-relevant footnotes or endnotes (Translation; acknowledgements for linguistic help; visiting affiliation and being on leave away from permanent base; we also think that source of funding falls into this category especially if the grant is being provided by an agency such as the Ministry of Defence). (Swales, 1985)

According to these criteria, Carmela Chateau Smith would have a minus score, ranking E (–5 to –12) on a scale from A (definitely NES) to E (definitely NNES). Far from invalidating Swales' somewhat ethnocentric criteria, this result raises several interesting questions. Do birth and education confer permanent NS status? Does the acquisition of additional languages (French, in the present case) actually modify linguistic ability and status? Does bilingualism lead to a higher threshold of tolerance for variation?

The less stringent criteria proposed by David Banks might produce more informative and useful results.

The articles were selected on the basis of two criteria:
(1) the article was written in English;
(2) at least one of the authors provided an address in an English-speaking country.
The object of this selection was to isolate those articles which were written by English speakers, so as to avoid possible interference from other languages. (Banks, 1991: 59)

It is possible that the presence of a co-author residing in an English-speaking country may help to avoid cross-linguistic contamination, but many factors enter into play in the publication of a scientific article, from the Instructions or Guidelines to Authors, through the process of peer-review (potentially by NNES), to the final stages of copy-editing.

Furthermore, most articles in the "hard sciences" (Weijen, 2012) are not the work of a single author, but are the fruit of teamwork, with several co-authors.

Many medical journals now ask the corresponding author to identify the roles and contribution of any co-authors, with some version of the following four criteria, as proposed by the International Committee of Medical Journal Editors:[5]

> Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
> Drafting the work or revising it critically for important intellectual content; AND
> Final approval of the version to be published; AND
> Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Unfortunately, none of the above criteria provides any information about the linguistic status of the contributors, and traces of copy-editing are difficult to identify, unless various versions of the article are available, as in the case of self-archiving of a pre-print copy, for example.

2.2 The Advice for Authors corpus

In order to explore the recommendations and advice available to researchers, a pilot Advice for Authors corpus was created, using *Sketch Engine* (Kilgarriff et al., 2014).[6] This corpus is composed of 27 articles providing advice to authors on how to write for publication, in the field of medical research; the total word count is 97,992 (or 119,185 tokens, including punctuation); the longest article contains 17,660 tokens, and the shortest contains 918 tokens. The corpus could, in fact, be considered to contain 20 articles (with an average word count of 6,000 tokens), since the shortest texts form eight parts of a series, with the same authors for each text. The articles were published over a ten-year period, between 2006 and 2015. Restricting the topic to a single field eliminates domain-specific variability, but no further precautions were taken to limit the range and scope of the language used. The authors (whether NES or NNES) are from

---

[5] Recommendations available online: (Accessed 3 February 2019)
<http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two>.
[6] Online corpus software and web service, with access to pre-loaded corpora:
<http://www.sketchengine.co.uk/>.

several different countries, of various scientific backgrounds, and offer advice on all stages of publication. This pilot corpus is therefore presented as a valid, random sample of the advice available to authors in various sub-fields of medical research. Brief details of the texts in the Advice corpus are given in Appendix I.

2.3 Exploring the corpus

Sketch Engine can be used to draw up word lists. It is possible to create word, lemma, or even lemmatized parts-of-speech lists, with an additional lowercase option. Unsurprisingly, in the basic frequency list for the Advice corpus (with the lowercase option), the fifteen most frequent words all belong to the closed set class, and are very short, confirming the Zipfian principle that: "as the relative frequency of a word increases, it tends to diminish in magnitude." (Zipf, 1935: 38). Among the first twenty words in order of frequency, only two are more than four letters in length, and both are lexical in nature, from the open-set class. These two words together, "journal" and "writing", encapsulate the main theme of the Advice corpus. The most frequent collocation for the lemma "journal/s" is not "writing" (22 occurrences), but "published" (43 occurrences). The implication is therefore that being published is twice as important as writing.

Writing for publication tends to be a choral work rather than a solo performance, even when an article has only one named author. In the Advice corpus, as expected, the plural forms, "authors" and "reviewers", found in 21 texts, are more frequent than the singular "author" and "reviewer". The word "review" occurs in 23 texts, underlining the importance of this task (Table 3).

Table 3: Advice for author.*, peer.*, and review.*

	T	F		T	F		T	F		T	F
author	25	136	peer	14	58	review	23	249	reviewer	16	87
authors	21	244	peer-review	5	5	reviews	10	65	reviewers	21	259
authorship	9	33	peer-reviewed	10	35	reviewed	12	26	reviewing	11	55

T= number of texts; F= number of occurrences

Focusing on the most frequent words common to all 27 texts in the Advice corpus can help to identify the main preoccupations of their authors, but most of the 29 results are short, closed-set words, with only two nouns: "journal" and "paper", again underlining the importance of publication.

The two most frequent words, consistent with most corpus results, are "the" and "of", exemplifying the most frequent frame in English: "the" *

"of" *, which produces a concordance of 740 occurrences, the most frequent of which are presented in Table 4, classified by frequency and by category of use.

Table 4: The most frequent frames and their uses.

Size and quality		Results		Options		Affiliation	
the quality of	40	the outcome of	12	the use of	11	the journal of	10
the number of	30	the results of	12	the choice of	6	the Departments of	8
the importance of	20	the end of	11	the style of	6	the University of	8
the lack of	8	the findings of	6	the type of	6	the work of	6
the accuracy of	7	the reporting of	6	the process of	5	the editor of	5
				the subject of	5	the names of	5
Semantics		Probability				the role of	5
the purpose of	6	the chances of	5				
the meaning of	5	the likelihood of	5				

The most frequent verb forms are, in order: is, be, are; have; should, will, may. This result provides useful information for NNES. It is possible to create concordances in *Sketch Engine* using Corpus Query Language (CQL). The following expression: [tag="VB.*"] [tag="VVN.*"] produces a list of 1157 uses of passive forms (any form of the verb "be", followed by any past participle). The large number of occurrences of "be" (429) suggests a further query, using the expression: [tag="MD.*"] [tag="VB.*"] [tag="VVN.*"] to find modal passive forms (250), presented in order of frequency: should (89), can (50), may (33), will (33), must (22), could (12), would (6), and might (5). The use of "should" is perfectly natural in a corpus of advice, but the proportion of "can" and "may" in relation to "must" suggests that few hard and fast rules can be proposed.

To examine this hypothesis, three CQL expressions[7] were used to identify the deliberate use of the "split infinitive", a well-known example of the prescriptive versus the descriptive approach to grammar (Yule, 2010: 85–87). The results (43 split infinitives, in 18 texts) indicate that splitting the infinitive is not a problem for most Advice corpus authors. The expression "adverb + to + verb" (66 occurrences) was used only once to deliberately avoid the split infinitive ("Do not be afraid respectfully to argue your case"). There were only eight cases where the expression "to + verb + adverb" (75 occurrences) could potentially have been replaced by the split infinitive.

---

[7] 1° "to" [tag="RB.*"] [tag="V.*"], 2° [tag="RB.*"] "to" [tag="V.*"], 3° "to" [tag="V.*"] [tag="RB.*"].