Article

An Exploration of the Sub-Register of Chemical Engineering Research Papers Published in English

Myrto-Panagiota Zacharof [1],* and **Anna Charalambidou [2]**

[1] Centre for Cytochrome P450 Biodiversity, Institute of Life Science, Swansea University Medical School, Swansea, Wales SA2 8PP, UK

[2] Media Department, Faculty of Arts and Creative Industries, Middlesex University London, The Burroughs, Hendon, London NW4 4BT, UK; a.charalambidou@mdx.ac.uk

* Correspondence: myrtozacharof1981@yahoo.com or m.zacharof@swansea.ac.uk; Tel.: +44(0)-7455858883

check for updates

Abstract: The increased pressures for high-volume, high-impact publications in English language and the high rejection rates of submitted manuscripts for publications present an often unsurpassable obstacle for (early career) researchers. At the same time, register variation of peer-reviewed journals—that can contribute to whether a paper is accepted for publication—has received little attention. This paper redresses this gap, by investigating the register (especially discourse moves and lexical choices) in 60 published, original-research articles on wastewater treatment published in four Chemical Engineering journals, with impact factor (IF) above 2. Our survey shows that chemical engineering research publications tend to comply with a set of requirements: multidisciplinarity, brevity, co-authorship, focus on the description of practical results (rather than methods), and awareness of non-specialised audiences. Lexical choices were analysed through frequency tables, phrase nets and word trees produced by data visualisation software (ManyEyes). It was found that less discipline-specific vocabulary is used in higher IF journals and this is interpreted within the current context of manuscript publication and consumption. This study concludes that data visualisation can provide an efficient and effective tool for prospective authors that wish to gauge telling details of the sub-register of a specific journal.

Keywords: chemical engineering; journal publications; lexical choices; word tree; phrase net; impact factor

1. Introduction

Chemical engineering science is a versatile, multifaceted field integrating physics, mathematics, biology and chemistry. Chemical engineers are employed both in the academic world and in a range of industries, from pharmaceutical and cosmetics fabrications, to hydrocarbons, food production and processing, and environmental pollution. In academia, in particular, the research activity occurring has factual outputs (such as communications of various character and nature) that are quantifiable; for instance, patents, presentations databases, protocols and publications [1]. Researchers, often non-native speakers, are expected to gather and produce information, process and evaluate it, and communicate it in a concise form [2,3]. The prevailing form of communication—and documentation—of research is scientific journal publication, and publishing in co-authorising teams is now the dominant *modus operandi* [1,4]. Researchers and scientists are under constant pressure to publish their results [5], as this would enhance their employment prospects and career development, their funding and consultancy prospects and their professional reputation [6,7]. Chemical engineering researchers are further challenged by the multifaceted nature of their discipline, since they are called to communicate their findings to a wider audience of fellow scientists, both during a manuscript's writing process and

its peer review. Moreover, high rejection rates of submitted manuscripts for publications have been observed, with 62% of published papers having been rejected at least once [8]. Numerous reasons influence rejection, including technicalities, such as limits in pages of publications per year (printed pages per issue, volume) and limited time between submission and publication [9], but mainly a lack of clear, succinct explanation of the findings and their significance to their scientific field [8], which is often attributed to the use of English language.

Although the acceptance of a manuscript for publication is an achievement, only high-quality publications in high-ranking scientific journals are widely accepted by the scientific community, authors' affiliations, and employment and funding bodies [10,11] For instance, the European Commission has formally recognised the importance of bibliometric indicators for policy purposes and is deeply engaged in and strongly encourages scientometric analysis [7]. The substantial number of predatory publishers [12,13], the increasing rate of generation of scientific findings, the globalisation of scientific communication through electronic media, the different sets of regulations regarding manuscript length, peer review and evaluation have increased the importance of determining and assessing the value of a publication. A publication is often evaluated through (a) the quality of the journal as reflected in its impact factor, and (b) the individual citations the article receives [6]. A journal's impact factor, despite being continuously and increasingly scrutinised [12], is the most popular numerical measure for the evaluation of a scientific publication.

The impact factor was originally designed as an aid to librarians all over the world, to select the journals that were most relevant to a library's users and aims [14,15]. It is a ratio calculated by the total number of citations a journal has received over the preceding two years divided by the total number of citations of articles published during that time [16–18]. Nowadays, impact factors have become a vital part of decision making regarding scientific impact [15], influencing decisions regarding career prospects, recruitment and appointments [1,17,18]. Therefore, understanding the specific requirements, including language and presentation issues, of journal writing is of utmost importance for every researcher [5]. Complying with stylistic requirements of the target journal would contribute to making a paper a good read, hence enhancing the potential of publication (by reducing editor's time) and increasing its citability potential [19].

Despite its importance, and even though some writing-related training across the curriculum at student level exists [4,20,21], training scientists in publishing their research findings is not an elemental part of chemical engineering education [19]. Post-doctoral researchers are expected to have already obtained the skills required for producing high quality publications during their postgraduate education or to have learnt by osmosis; for instance, through reading published manuscripts from fellow researchers, a tactic that might be highly ineffective and time consuming, and can lead to failed attempts to publish [22]. Several books have been published offering guidelines for writing papers [23] in science, chemistry and engineering. However, these books give general advice on the structure the papers need to have related to the analysis of experimental data of quantitatve and qualitative nature without focusing on the use of language [24]. In addition, there is concrete evidence of lexical variation of texts within the same academic discipline, depending on the type of publication (i.e., journal article, research proposal, scientific poster, textbook, popular science article) and, consequently, on its intended audience (expert, scientific, student, general public) [21,25–27]. The issue of register variation among articles published in different types of peer-reviewed periodicals has received little attention and is a much-needed addition to chemical engineering education at university level, and researcher development, in general.

Hence, this paper aims to identify and investigate the register of 60 published manuscripts in four different journals of impact factor above two. All four journals publish original research papers in the field of chemical engineering science, and specifically in one of its most prominent and complex subject areas, environment conservation and sustainability, focusing on wastewater treatment (Figure 1). This study explores possible links (or lack thereof) between the impact factor and scope of each journal, on the one hand, and the register of the manuscripts (with a focus on lexical choices and

discourse moves), on the other. To the authors' best knowledge, register variation between different types of *published, professional original research* articles has not been researched. This paper thus aims to investigate how lexical choices and the content of scientific manuscripts relate to the advertised scope and impact factor of the journal, in which they are published. This can contribute in helping chemical engineering researchers better adapt their papers to suit the specific register of their chosen journal, so as to positively influence their publication record and career prospects, and to attract citations and possible collaborations.
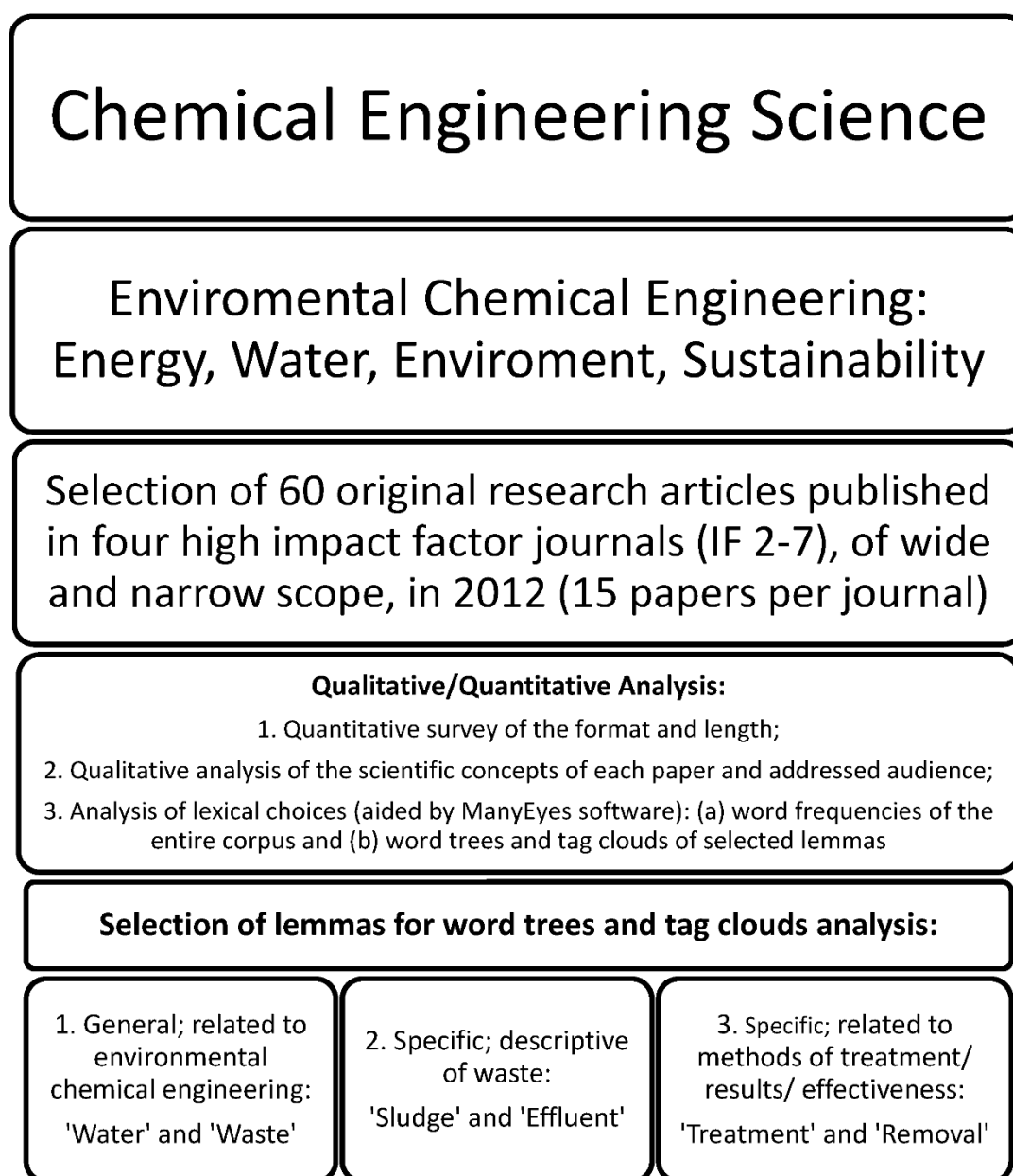
Chemical Engineering Science

Enviromental Chemical Engineering: Energy, Water, Enviroment, Sustainability

Selection of 60 original research articles published in four high impact factor journals (IF 2-7), of wide and narrow scope, in 2012 (15 papers per journal)

Qualitative/Quantitative Analysis:

1. Quantitative survey of the format and length;

2. Qualitative analysis of the scientific concepts of each paper and addressed audience;

3. Analysis of lexical choices (aided by ManyEyes software): (a) word frequencies of the entire corpus and (b) word trees and tag clouds of selected lemmas

Selection of lemmas for word trees and tag clouds analysis:

| 1. General; related to environmental chemical engineering:

'Water' and 'Waste' | 2. Specific; descriptive of waste:

'Sludge' and 'Effluent' | 3. Specific; related to methods of treatment/ results/ effectiveness:

'Treatment' and 'Removal' |

Figure 1. Schematic representation of the methodology developed and followed in this case study.

2. Materials and Methods

2.1. Materials

Four journals related to chemical engineering, with an impact factor above two, were selected, namely *Water Research* (**WR**), *Waste Management* (**WM**), *Chemical Engineering Science* (**CES**) and *Chemical Engineering Journal* (**CEJ**). The selection of the four journals was based on the following criteria:

1. The topics the journal addresses, as advertised in the website of each journal, needed to include environmental chemical engineering, wastewater treatment and management;
2. The intended scientific audience, as advertised on the journal's website, needed to include chemical engineering professionals;
3. The journal needed to have at least 15 original research articles published in 2012 focusing primarily on various aspects of wastewater treatment and management, for example industrial and agricultural wastewater, separation science, etc.;
4. The journal's impact factor needed to be above 2 [28]

CES and **CEJ** were considered journals of wider scope, due to the fact that they accommodate publications on a range of subject areas within chemical engineering (Table 1), while **WR** and **WM** were regarded as journals of specialised scope, due to their more concentrated focus on areas relevant to environmental chemical engineering. All four journals publish various types of papers related to environmental chemical engineering and its major areas of energy, water, environmental impact and sustainability (Table 1). Fifteen original papers, i.e., research-related scientific manuscripts describing, analysing and discussing experimental trials and case studies were selected from each journal, totalling sixty papers, all of which were published in 2012. We focused only on a single year of publications to provide a multi-layered, focused investigation, controlling, as much as possible, additional variables (such as year of publication) that might also have affected the register of the papers. As language trends and all other requirements of journal writing evolve, confining the dataset to a single year would eliminate issues of diachronic variation. We chose to focus on volumes published in 2012, as that would give us a period of five years to track accumulated citations. We limited our enquiry to original research articles about waste-water management to have a comparable sample. On the whole, restraints in the type of papers selected were placed to ensure a homogenous, consistent sample, in order to extract meaningful results and draw useful conclusions, since the vast majority of published papers in sciences, including chemical engineering, correspond to the type of factual research-related manuscripts. By looking at journals of both wider chemical engineering and narrower (water/waste) focus, we aim to examine whether the focus of the journal, rather than the focus of the article, has an observable effect on the register of each paper.

Table 1. Summary description of the scope of each journal as published on the journals' websites.

Journal	Water Research	Waste Management	Chemical Engineering Science	Chemical Engineering Journal
Affiliations	International Water Association (IWA)	-	-	-
Website	www.journals.elsevier.com/water-research	http://www.journals.elsevier.com/waste-management/	http://www.journals.elsevier.com/chemical-engineering-science/	http://www.journals.elsevier.com/chemical-engineering-journal/
Publisher	*Elsevier B.V.*			
Audience	Chemists, biologists, microbiologists, immunologists, limnologists, civil engineers, sanitary engineers and chemical engineers.	Scientists, engineers and technical managers concerned with waste treatment and the engineering problems related to environmental protection laws. Scientists, engineers, and managers, regardless of their discipline, who are involved in scientific, technical and other issues related to solid waste management.	Industrial and academic researchers in chemical and process engineering.	Chemical and process engineers, applied chemists and product engineers, biochemical engineers and biotechnologists
Impact Factor (IF) [1]	6.942	4.030	2.895	6.216
Publication Rate	20 issues per year (1 volume per year)	10 issues per year (1 volume per year)	12 volumes per year (1 issue per volume)	No issues, 39 volumes per year
Mean Number of publications per issue/volume	36	25	20	56
Types of papers published	Full papers, review papers, comments	Full papers, review papers, letters to the editor, columns	Original papers, review articles, short communications, letters to editors	Original papers, review articles, short communications, letters to editors
Scientific subjects published	No specific scientific sections, the journal interested in water quality and its management. It publishes original research on treatment processes for municipal, agricultural and industrial water and wastewaters, water quality standards and analysis by chemical, physical and biological methods	Emphasis is placed on integrated approaches, major areas in which papers are solicited: generation and characterization, minimization, recycling and reuse, storage, collection, transport, and transfer, treatment (mechanical, biological, chemical, thermal, other), landfill disposal (including design, monitoring, remediation of old sites), environmental considerations, financial and marketing aspects, policy and regulations, education and training, planning and implementation.	Publication of papers on the fundamentals of chemical engineering, including. Industrial areas covered by the journal include biotechnology, chemicals, energy, food, materials, microelectronics, nanotechnology, specialty chemicals and pharmaceuticals. biomolecular and biological engineering, biochemical and bioprocess engineering, energy, water, environment, and sustainability materials engineering, particle technology; process systems engineering reactions, separations science and technology	Three aspects of chemical engineering: chemical reaction engineering, environmental chemical engineering, and materials synthesis and processing.
Abstract	Concise and factual, descriptive (up to 250 words)			
Graphical abstract	Optional			
Highlights [2]	Mandatory			

[1] 2012 Journal citations report by Thomson Reuters http://thomsonreuters.com/journal-citation-reports/ (last accessed 29 March 18).

[2] Highlights are a short collection of bullet points that convey the core findings and provide readers with a quick textual overview of the article. These three to five bullet points describe the essence of the research (e.g., results or conclusions) and highlight what is distinctive about it. There should be a maximum of 85 characters, including spaces, per highlight.

2.2. Methods

A multi-layered analysis of the collected papers was devised, employing a mixture of qualitative and quantitative methods as well as lexical analysis methods (Figure 1). Quantitative analysis related to the length of the papers (word counts, number of authors, references, pages, tables, and figures) and was conducted in order to identify similarities and common trends, using Portable Document Format (.pdf) to MS Office Word 2007 converter software by free pdf solutions (www.freepdfsolutions.com). Further analysis was done with MS Office Excel 2007, using linear regression analysis to obtain the average data and estimate standard error and standard deviation (below <5%).

Qualitative analysis of the corpus followed previously published methodologies [29,30] focusing on the main scientific concepts each published manuscript was addressing. Each paper was broken into clusters according to the classic practical sciences report writing style, which is introduction, materials and methods, results, discussion and conclusion. Each paragraph contained in the clusters was then conceptually analysed, aiming at a literal description, analysis and understanding of the stated research including its methodology, findings, conclusions and addressed audience. Two independent examinations were carried out by each of the authors to minimise each reader's subjectivity and bias regarding the manuscript content. Any disagreements that arose were resolved after thorough discussion between the authors, until a unanimous consensus was reached.

Lexical variation was examined through computational analysis of word association and frequencies, facilitated by ManyEyes software (www-958.ibm.com). ManyEyes visualisation, which pioneered easy and free visualisation of data, is a very effective resource, as the user is not required to have technical expertise, or to buy and/or download a programme or application. This software allows the creation of visualisations from large datasets. The following three visualisations were chosen, as they were the most pertinent to the type of data (text) and research objectives (see supplementary material Figures S1–S3:

- **Tag clouds**: Visualisations of word frequencies, which enable the researcher to see how frequently a given word appears in the corpus.
- **Phrase nets**: This visualisation shows patterns of frequent pairs of words. Words are connected when they are separated by 'and'; 'of the'; 'is', space, 'at', 'a', 'is', and 'the' in the source text.
- **Word trees**: This visualisation enables the analyst to pick a word or phrase and shows all the different contexts (i.e., immediately prior or upcoming text) in which the word or phrase appears.

The ManyEyes software can account for large amounts of text and provide accurate and fast calculations, reducing the researcher's bias. It can highlight the contrast between our intuitions about word use and actual patterns in authentic language. An additional benefit is that it has the potential to find exceptional cases. Analytical tools from corpus linguistics were employed for the analysis and interpretation of word frequencies and word association in the various journals [31,32].

3. Results

In order to better contextualise the findings of the fine-grained analysis of the lexical choices in the different journals, an overview of the format and length of the papers is provided, followed by qualitative analysis of their target audience.

3.1. Format of the Collected Papers

In practical sciences, including engineering, manuscripts are generally considered shorter in length compared to liberal sciences and arts [33]. Commonly within a breadth of 6 to 12 printed two-column pages, including tables, figures and references [34] the authors are expected to present and discuss their study. Reduction in volume and size of research papers has been widely implemented to physical sciences journals, due to the constantly increasing rates of submission and the ensuing need to accommodate a higher number of published papers within a journal's printed issue or volume [35].

The shorter length of such papers is also enabled by the possibility to visualise findings in meaningful figures that need little or no explanation, as well as tabulating the core findings [36–38].

These findings are also supported in this case study. The papers' lengths were between 8 and 13 printed pages, including figures and tables, with word counts between 6800 and 9700 words including references, highlights, abstracts, tables and figures (Figure 2). With regard to the length of the papers, similarities were found between **CES** and **WR** (average 11 pages, 8300 words) and **CEJ** and **WM** (average 8 pages, 6700 words). References in all papers ranged between 36 and 48, with similar trends found between the wider scope journals **CES** and **CEJ** (on average 37 references) and the more specialised scope **WM** and **WR** (on average 41 references) (Figure 2).

Figures and tables are the core part of the published manuscripts, varying in number from 4 to 10 figures and 3 to 5 tables. This is also reflected in the analysis of word frequencies, where the words "table" and "figures" are among the top ten most frequently used words among all the journals (Table 2), implying that the text's primary function, especially in the results sections, is to comment upon the visual parts of the papers. Within that context, the text serves for analysing, explaining and discussing these visual aids to the audience.

The quantitative analysis suggests that the selected papers from each of the four journals share similar quantitative characteristics, thus rendering the four datasets comparable. Also, there does not seem to be a link between the impact factor or scope of each journal and the volumetric characteristics and layout of the published articles.
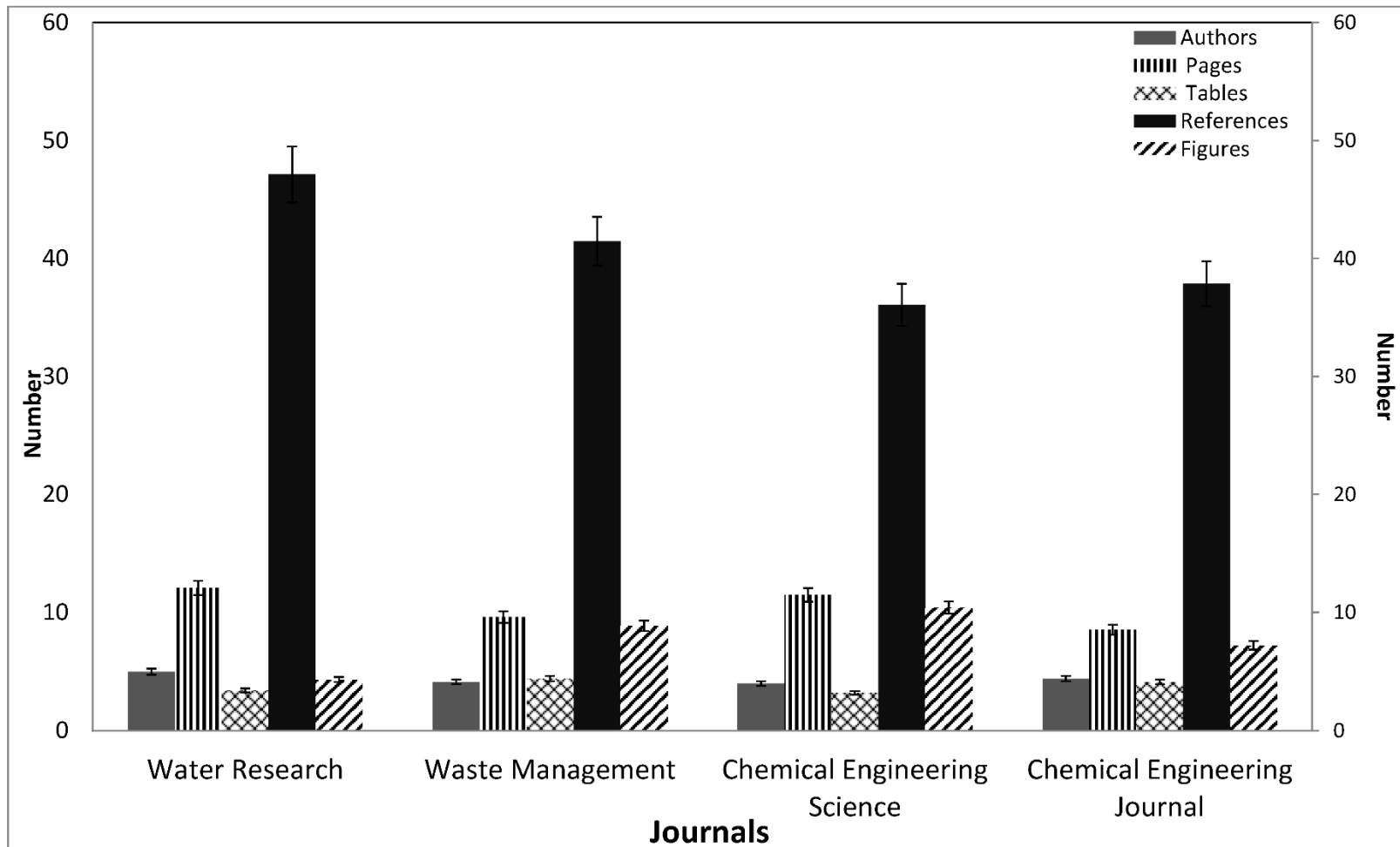
Figure 2. Volumetric characteristics of the analysed published papers, including number of authors, pages, tables, references and figures per journal.

Table 2. The top ten words occurring in each journal and their total number of occurrence.

Journals							
Water Research		Waste Management		Chemical Engineering Journal		Chemical Engineering Science	
Words	Amount	Words	Amount	Words	Amount	Words	Amount
water	700	**waste(s)**	1151	**concentration(s)**	602	**water**	478
treatment	540	leachate	748	model	470	**pH**	405
concentration(s)	520	landlfill	671	**fig**	468	**concentration(s)**	351
effluent(s)	520	lysimeter	503	gas	382	removal	326
removal	410	**fig**	396	mm	327	**fig**	325
samples	408	emissions	376	CO_2	323	**wastewater**	296
wastewater	350	system	332	rate	322	**mg**	283
environmental	276	collection	298	absorption	283	**treatment**	277
mg	307	**treatment**	297	reaction	267	min	244
table	264	cod (chemical oxygen demand)	289	**pH** phase	262	phosphate	239
Total number (15 papers/journal)	144,798	115,491		101,519		125,910	

3.2. Multidisciplinary Nature of the Analysed Papers

Despite their moderate size, all published manuscripts were the outcome of collaborative efforts, with the mean number of authors being four. The multidisciplinary nature of chemical engineering calls for extensive cooperation, since specialists from many disciplines are required to perform the integral experimental trials and analysis [39–41]. The diverse nature of the published papers in chemical engineering was clearly reflected in this study, by the subject category (Figure 3) and audience distribution (Figure 4).

Out of the 60 papers investigated, waste treatment and management was approached from a range of perspectives, from biochemical engineering to environmental chemistry, to mechanical, electrical or civil engineering (Figure 3). Interestingly, the advertised scope of each journal is broader compared to the categories that emerged from this research.

The topics covered in the papers in our sample were found to be of potential relevance to a broad audience, not restricted to academia, but also to other bodies such as policy regulators, small- and medium-size companies and enterprises, or environmental agencies. In fact, the collected papers addressed an audience of thirteen categories varying from water and environment specialists to microbiologists and chemists, as well as governmental bodies, water and wastewater companies (national, private) or regulatory policies agencies and law developing and forming bodies (Figure 4). This is a divergence from the advertised audience in the website of each journal, where the focus is on specialists in chemical engineering audience within the field.

The content analysis of the papers showed that in **CEJ** and **CES** there was a stronger tendency, compared to **WR** and **WM**, to appeal to the industry. This could be attributed to the nature of studies, i.e., dealing with trials at a pilot plant scale with large volumes of materials. This type of studies is more attractive to the industry, since the authors have not only proven their concept but have also applied it to a large scale. In contrast, **WR** and **WM** are primarily addressing an academic audience, with **WM** publishing also on topics that are of interest to the regulatory authorities of each country and globally regarding waste. Thus, patterns emerged about the nuances of the addressed audience in the published manuscripts of each journal that were not clearly communicated on the journals' websites.

The wide range of potential audiences of the published papers emphasises the need for clear, concise and easily understood language, as readers coming from different academic disciplines, even in close proximity, might fail to comprehend the concepts and rationale expressed in the manuscripts.
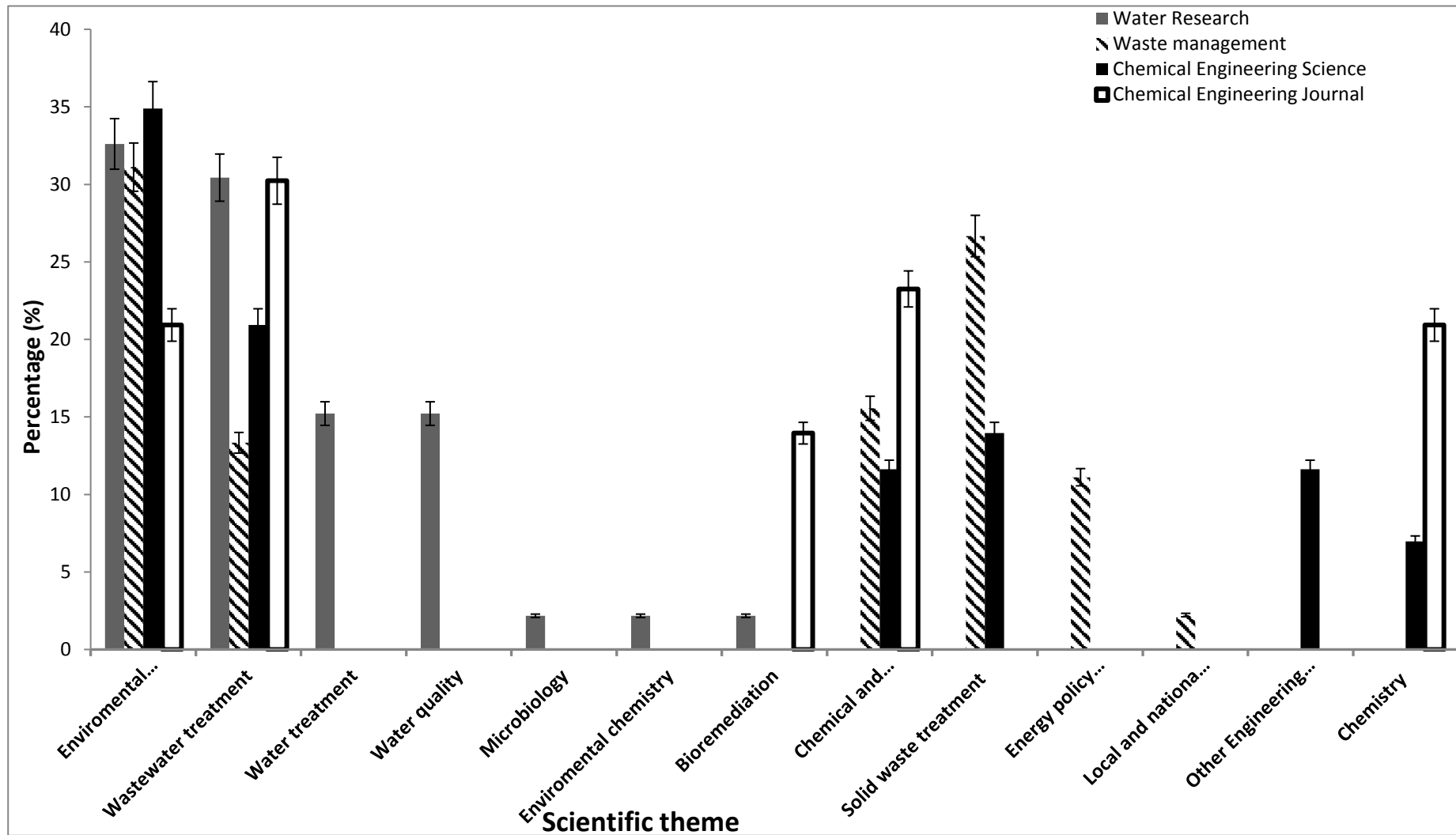
Figure 3. Scientific theme distribution of the analysed published manuscripts.
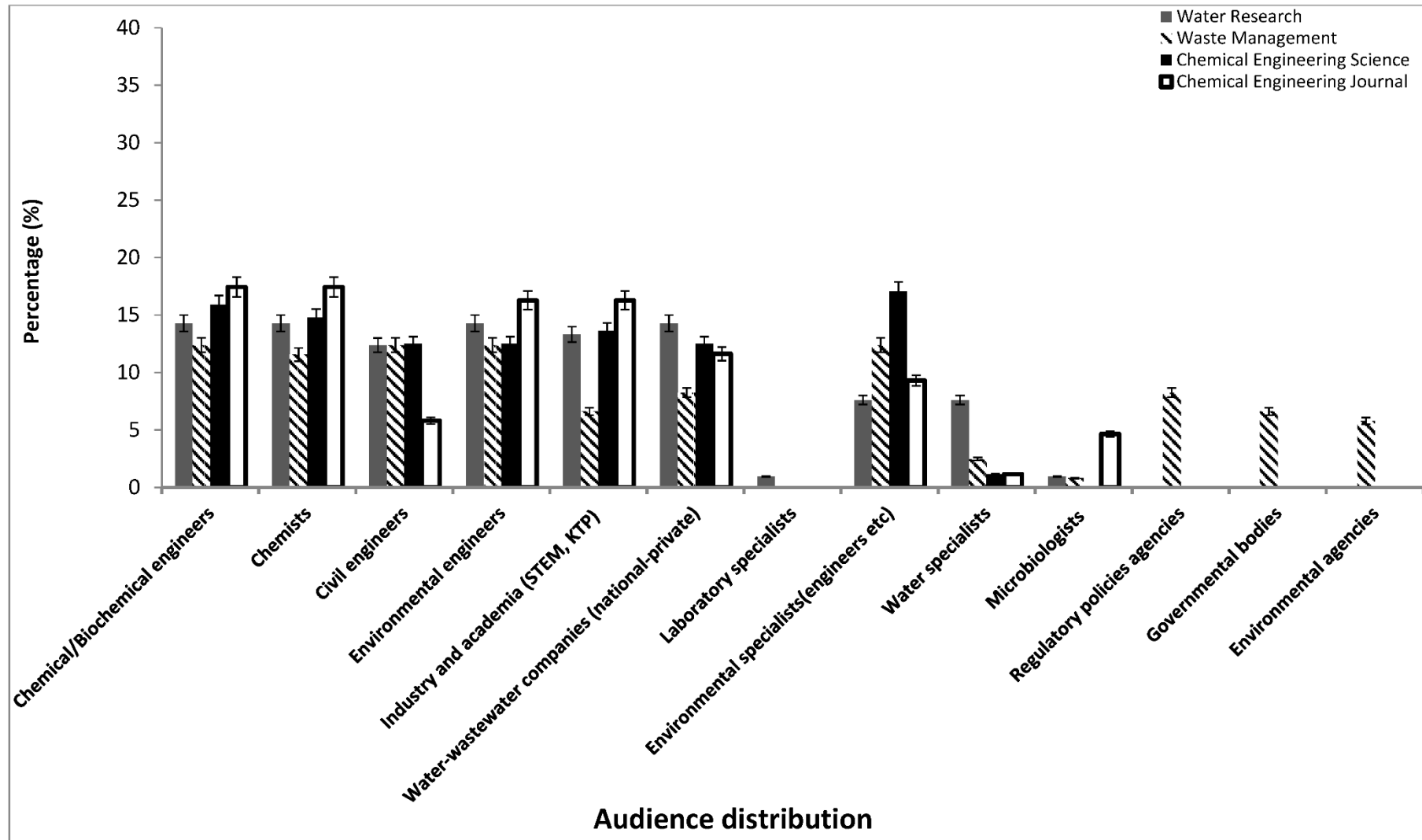
Figure 4. Distribution of potential types of audience of the analysed published manuscripts.

3.3. Analysis of Lexical Choices

As suggested in the introduction, lexical choices are intimately linked to the text type and intended audience. In order to investigate lexical variation among different types of journal articles on wastewater treatment and management, the most frequently used words in the corpus were analysed, as well as the context of the occurrence of certain key words, and correlations were explored between the results and the type of journal (wider or specialised scope) and the journal's IF. Six lemmas were chosen, to explore the context in which certain key terms are employed and variation in the specific meaning that is ascribed to them (Figure 1). These terms include *water* and *waste,* which are generally used when referring to the environment and would be expectably mentioned mostly in the introduction and discussion or conclusion parts of the papers; two lemmas specifically related to and descriptive of waste, *sludge* and *effluent*, which could be found throughout the manuscript and especially in the results section; and, finally, two lemmas related to the experimental methodology used and the achieved results and relevant conclusions, *treatment* and *removal* (see Figure 5 for the frequency of occurrence of these six lemmas in each journal). The co-text of the lemmas and consequently the specific meaning they accrue because of their context of use (context is taken here as immediately prior and upcoming text, see [42]) were analysed based on 'word trees' and 'phrase net' visualisations (see supplementary material Figures S1–S3). Below, the key findings of the analysis through the ManyEyes phrase net and word tree visualisations of the six lemmas are outlined.

In **CEJ**, the lemma "water" was found 546 times in a total of 15 papers and, as the analysis of the word tree visualisations showed, was mainly conceptualised as a resource (ground water, surface water, wastewater); either potable or as liquid waste. Focus was placed on reuse (removal of harmful elements and use as washing water), recycling (water reclamation in the scope of cost reduction, environmental load) and treatment (removal of toxic metals such as lead, copper, harmful substances i.e., pesticides, hormones, pharmaceuticals) of water, focusing on wastewater treatment.

In the 15 **CES** collected papers, "water" occurred 176 times and was mainly understood as a tool within the context of a chemical reaction, water as an aid in a chemical process for example in the form of steam during sterilisation, as solvent, as treatment method for other elements or as a component to other substances.

As regards the 15 **WR** and 15 **WM** papers, "water" was found 792 and 244 times, respectively, and, as the word trees suggest, it was conceptualised as a matter worthy of research, a resource, an object of analysis regarding quality, safety, treatment (potable water treatment, i.e., softening, salts and metals removal), wastewater (liquid waste of industrial, municipal, domestic, agricultural, slaughterhouse, food, tanning industry origin), a resource, and water cycle (water as an environmental resource, ponds, rivers, lakes).

The word "waste" was found 413 times in **CEJ**, and was used to indicate a problematic material that has to be treated, managed and disposed. It was commonly found immediately preceding the term 'water', forming the compound "wastewater", referring to the liquid or semi-liquid, semi-solid nature of waste.

Similar to the use of the lemma "water", "waste", occurring 131 times in **CES**, was mainly conceptualised as part of or a tool for a chemical reaction, a part a chemical process, the substrate or sample where the chemical process is applied on, as a component to other substances.

In **WR** and **WM**, "waste" was found 462 and 1150 times, respectively, and, similar to the word "water", it was used in the context of a research subject deriving of numerous sources, a subject of analysis regarding quality and treatment, but as well as a component or a resource for the production of other materials.
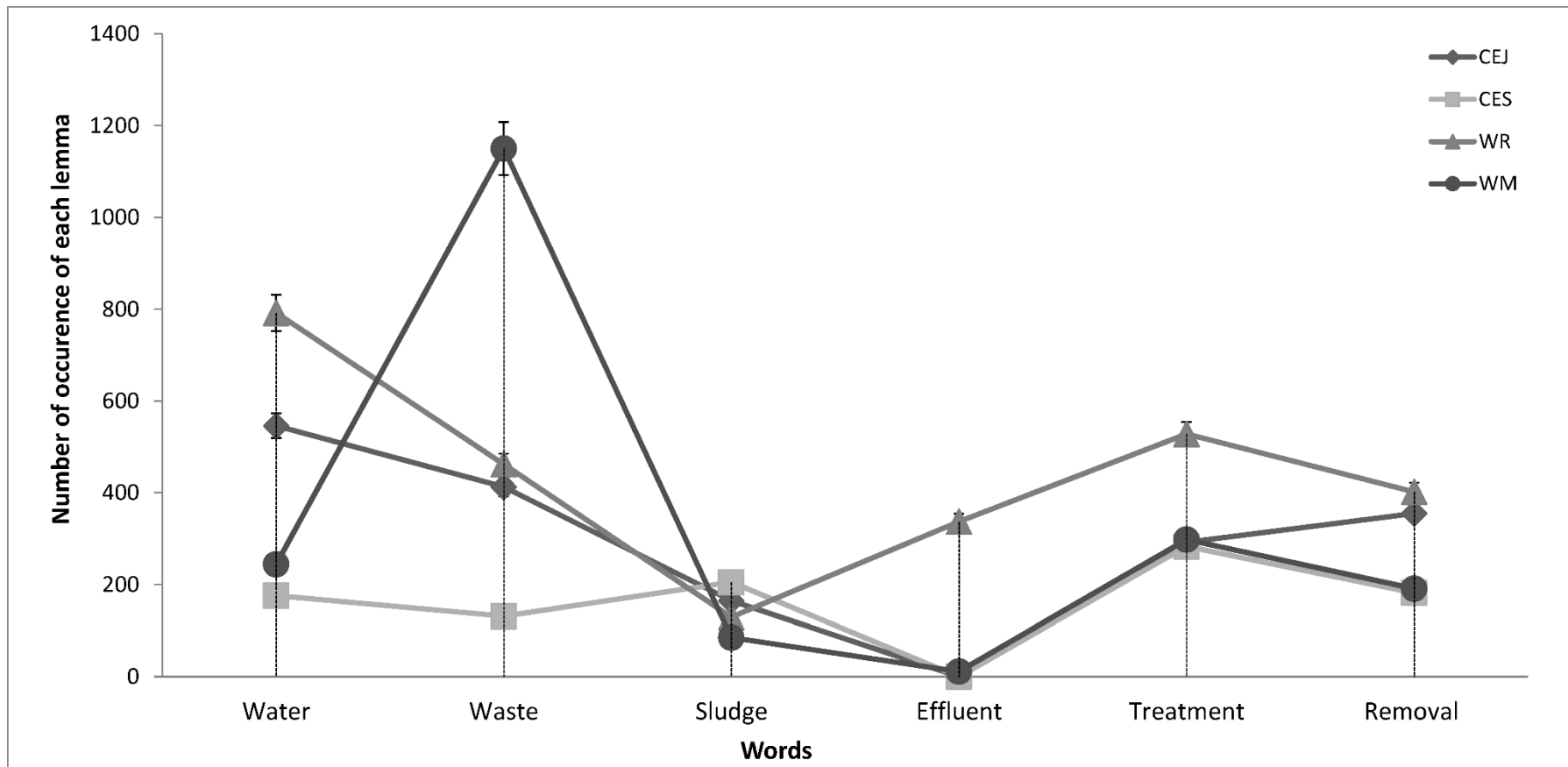
Figure 5. Distribution of the selected keywords among the analysed manuscripts.

Both lemmas "water" and "waste" were routinely found in the manuscripts of each journal, and they were among the top 10 most frequently encountered words in the manuscripts, and were used in high frequency either combined—i.e., wastewater—or separately (Table 2). However, as the analysis of the visualisations indicates, in **CEJ** and **CES**, the terms were recurrently employed in contexts different from **WR** and **WM**. In **CEJ** and **CES** the words are used in a rather specialised context compared to **WR** and **WM**, an interesting observation that did not confirm the authors' expectations, since both journals are of wider scope (Table 1), and it was anticipated that a less restrictive use of the term would have been encountered. In **CEJ** and **WR**, the words were found in analogous amounts; while in **CES** the amount of use was very limited, suggesting the use of a scientific specialised vocabulary (e.g., the terms "liquid" or "fluid" or "solvent", were preferred over "water"). On the other hand, **WM** stood out, since the lemma "waste" is used very frequently, suggesting a broader approach to the subject (i.e., industrial, agricultural, slaughterhouse, domestic, municipal waste).

Further investigation of the observed trends was achieved by examining the word associations for "sludge", "effluent", "treatment" and "removal" (Figure 1), as can be deduced from the visualisations.

In **CEJ** the word "sludge" was found 165 times and was referred to as a problematic, potentially harmful and hazardous material coming from waste. On the other hand, in **CES** "sludge", found 205 times, was used to describe a muddy, murky, highly viscous thick material in need of processing or treatment, which was not necessarily harmful or indicative of a problem. In **WR** and **WM** "sludge" occurred 129 and 85 times, respectively, and had a far more complex meaning, as it was used in the context of harmful material coming out of waste, physically looking like a murky, muddy, soil-based material, liquor or concentrated liquid of a semisolid nature coming out of the process of treating sludge.

In **CEJ** and **CES** the word "effluent" was not found, implying the absence of the mention of any mechanical treatment process that would separate the solid from the liquid phase of sludge, such as filtration, and the absence of any treatment involving large scale processes, a finding that is related with the subject and audience distribution of the journal as defined by the journal's author guidelines. In **WR** the word "effluent" was found 337 times, and was used to explain any liquid coming out of a waste treatment or of waste producing process, while in **WM** it occurred only eleven times, and was used for any liquid discharged of a leaching related process.

In **CES** the word "treatment", occurring 283 times, referred to any method and/or process used to uncouple sludge or wastewater of its harmful, dangerous, hazardous, toxic elements. In **CEJ** and **WM**, "treatment", found 292 and 298 times, respectively, was used to describe any process used, developed or applied to water, wastewater and sludge, without specifically explaining whether it is done to remove hazardous substances or simply for treatment. In **WR** "treatment" occurred 528 times and had a more generic meaning, referring to any process in which waste is involved, for example anaerobic digestion for combined heat and electricity production, to technologies or systems used to remove the harmful components.

Finally, the word "removal", found 325 and 182 times in **CEJ** and **CES** respectively, was employed to refer to any method and or process used to recover nutrients from the waste or remove all the components that are harmful and/or toxic. In **CES** "removal" also represented the main scope of the project developed in the article. In **WR** and **WM**, "removal", occurring 402 and 191 times, was used in the context of referring to any process or method applied to the removal of harmful elements from the discharged effluents, wastes, sludge or wastewater.

4. Discussion

The analysis of word associations and word frequencies (see supplementary material, Figures S1–S3) showed emerging trends in lexical choices that also have implications about the specific subject area and approach preferred in each journal, and—interestingly—do not necessarily correspond with the advertised scope of each journal. Among the four journals selected, **WR** and **WM** were considered of specialised scope (based on their advertised scope), and were thus expected

to have a highly specialised and technical lexis, whereas **CEJ** and **CES** were expected to use less discipline-specific lexis, due to their wider range of scientific areas and potential audiences, as described on the journals' websites (Table 1).

However, these expectations were not completely supported by the findings. **WM** and **CES** were found to be the journals where more specialised vocabulary is used, especially in **WM**. The high occurrence of discipline-specific vocabulary was not only associated with the scale of the experiments, but also with the emphasis on the methodology and experimental phase (meaning the size, the accommodating volume of the equipment and the size of volume eligible to be processed by the proposed methodology) rather than the results and their impact and applicability. The technical vocabulary was mainly associated with quantifiable data, experimental trial chemical reaction and processing, for example "model", "fig.", "lysimeter", "system", "reaction", as the lists of the most frequently used words in these journals indicate (Table 2). This finding is also supported by the close reading of the published manuscripts, which revealed that the manuscripts in **WM** and **CES** refer to specialised and complex methods of chemical engineering (Figures 3 and 4). For example, in **CES**, instead of plainly using "water", other related terms are used, such as "concentrations" and "phase", which point towards chemical processing, whereas in **WM** terms related to water such as "leachate" are used to point towards residuals of solid wastes.

On the other hand, CEJ and WR use a less discipline-specific vocabulary, with salience of terms "wastewater", "effluents", and "samples", which are far less frequent in **CES** and **WM** (Table 2). The findings from the qualitative analysis of **CEJ**, that show the multidisciplinary nature of the papers, coincide with the results of the lexical analysis. This further confirms that the use of a less discipline-specific vocabulary enhances the readability of the journal, beyond narrow disciplinary confines, and can thus address a range of audiences, including industry and policy regulators. On the other hand, while **WR** was found in the qualitative analysis to target in the main an academic audience, the use of simplified vocabulary boosts its readability by scientists from a wide range of disciplines. The analysed **WR** papers index a holistic approach to water-related research, focusing on the findings of the experimental trials and their applicability to the society, addressing social, financial and legal aspects. This also corresponds with the frequent use of the lemmas "removal" and "environmental".

Such findings indicate a correlation between increased intelligibility (beyond the narrowly conceived discipline of environmental chemical engineering) and citability of the journals, since **WR** and **CEJ** have the highest impact factors, at 6.942 and 6.216, respectively. Technology has facilitated a major shift in knowledge exchange from exclusively printed media to a combination of available online, easily downloadable articles and printed media, expanding significantly the visibility of a paper, as the readers are not relying only on the printed resources that exist in libraries and repositories across the world [18,38].

Literature searches are not necessarily guided by advisors, supervisors or assisting librarians, and are being partially replaced by specialised research engines such as Google Scholar or Scopus and the relevant webpages of the main academic publishers such as Springer, Elsevier, Sage or Wiley. This leads to the reading of the majority of published papers, on an individual unsupervised basis, by an audience that may not have an extensive knowledge of the subject (postgraduate students and early career professionals, researchers, academics and fellows), who may be novices on the specific subject area of the article. Employing highly complex, scientific lexis might not facilitate the understanding of the manuscripts by readers and will possibly result in lower citability. This could explain the association that was found in this study between more accessible, less specialised vocabulary and higher IF.

When comparing these findings to the advertised scope of each journal, certain differences were found. Among all four journals, only **WR** published papers that reflected the journal's very broad approach, focusing on innovation without disregarding new approaches to current techniques. **CEJ** and **CES** had a narrower thematology, addressing highly specific subjects contrary to the journals' advertised spectra. In the published manuscripts, emphasis was placed on optimisation of existing

methods, mainly chemical treatments, rather than innovation, which cannot be applied as easily and quickly. A similar tendency was found in **WM**, where, in spite of the wide advertised array of publishing subjects, the published manuscripts did not cover such a wide spectrum, and focused primarily on waste management and relevant regulations, reflecting the anisomorphy between the advertised and the actual scope of the journals.

5. Conclusions

This is a case study and results are not unproblematically generalisable across journals of practical sciences, let alone all disciplines. As our study has shown, there are journal-specific trends, even within the same narrow discipline, and of course there is significant register variation among disciplines as regards expectations of organisation, discourse moves, presentation of materials, and lexical choices. In addition, journal articles, as with any other text type, are constantly changing and are shaped by wider contextual factors, and therefore characteristics of their register are also evolving.

Nevertheless, this snapshot of trends in published chemical engineering research offers an insight into the register of journal articles and the implications of publishing research findings that can be extended beyond the four journals. Some tentative conclusions that could be deduced regarding the lexical and thematic choices in original chemical engineering research articles and which could be incorporated in learning and teaching material for chemical engineers, but also researchers from other disciplines that seek to publish their research, include the following.

Highly discipline-specific vocabulary use, including extensive use of acronyms, should be avoided where possible, to aid favourable consideration of manuscripts at higher IF journals and to increase the citability potential of the article.

There is a complex relationship between the thematology, the audience and the scope, as they are advertised in the journal's website, and the actual published manuscripts. The fact that journal guidelines are not foolproof representations of a journal's actual remit of publications is not a novel finding. What our research shows is that data visualisation can provide a quick way for researchers to assess the specific areas that are most likely to be published in the journal. More specifically, visualisations of word choices and associations, which can be easily and quickly done with the aid of freely available software, are a very powerful tool in providing an accurate overview of both the preferred content and approach of each journal, as well as its preference as regards to lexical choices.

We used ManyEyes, a free and user-friendly solution for novice academics with no programming expertise, who wish to get an overview of the scope of their target journal. Alternative visualisation tools as they become available can also be used, on the same basis that we describe in our paper. They can be an indispensable tool for chemical engineering students and novice researchers that wish to gain an emic understanding of the actual scope of the plethora of journals within each discipline, without having to engage in the labour-intensive close reading of a large corpus of published papers.

It would be desirable to extend this research to similar investigations to journal articles of a range of topics, disciplines and periods of time, so that the findings can be more representative of journal register in English. Further exploration of links between lexical choices and citability, impact factor, new media use and altmetrics (online traffic of journal's published manuscripts) could lead to the development of a methodology that would help researchers better understand the specific register of their target journal.

References

1. Abramo, G.; D'Angelo, C.A.; Rosati, F. Career advancement and scientific performance in universities. *Scientometrics* **2014**, *98*, 891–907. [CrossRef]
2. Massoudi, M. Can scientific writing be creative? *J. Sci. Educ. Technol.* **2003**, *12*, 115–128. [CrossRef]
3. Vanclay, J. Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics* **2012**, *92*, 211–238. [CrossRef]
4. Beall, H. Expanding the scope of writing in chemical education. *J. Sci. Educ. Technol.* **1998**, *7*, 259–270. [CrossRef]
5. Uskokovic, V. Ten Commandments for writing a meritable scientific paper. *J. Postdr. Aff.* **2012**, *2*, 2–7.
6. Bosquet, C.; Combes, P.-P. Are academics who publish more also more cited? Individual determinants of publication and citation records. *Scientometrics* **2013**, *97*, 831–857. [CrossRef]
7. Porter, A.; Rafols, I. Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics* **2009**, *81*, 719–745. [CrossRef]
8. Woolley, K.; Barron, P. Handling manuscript rejection insights from evidence and experience. *Chest* **2009**, *135*, 573–575. [CrossRef] [PubMed]
9. Drummond, C.W.E.; Reeves, D. Reduced time to publication and increased rejection rate. *J. Antimicrob. Chemother.* **2005**, *55*, 815–816. [CrossRef]
10. Barrow, L.H. Searching for educational technology faculty. *J. Sci. Educ. Technol.* **2003**, *12*, 143–147. [CrossRef]
11. Tscharntke, T.; Hochberg, M.E.; Rand, T.A.; Resh, V.H.; Krauss, J. Author sequence and credit for contributions in multiauthored publications. *PLoS Biol.* **2007**, *5*, 13–14. [CrossRef] [PubMed]
12. Beall, J. Beall's List of Predatory Publishers 2013. Available online: http://scholarlyoa.com/publishers (accessed on 29 March 2018).
13. Moore, A. The garbage collectors—Could a particular sector of author-pays journals become silently acknowledged collectors of scientific waste? *BioEssays* **2009**, *31*, 821. [CrossRef] [PubMed]
14. Amin, M.; Mabe, M. Impact factor: Use and abuse. *Int. J. Environ. Sci. Technol.* **2000**, *1*, 1–6.
15. Fassoulaki, A.; Sarantopoulos, C.; Papilas, K.; Patris, K.; Melemeni, A. Academic anaesthesiologists' views on the importance of the impact factor of scientific journals: A North American and European survey. *Can. J. Anaesth.* **2001**, *48*, 953–957. [CrossRef] [PubMed]
16. Ovalle-Peradones, M.-A.; Gorraiz, J.; Wieland, M.; Gumpenberger, C.; Olmeda-Gomez, C. The influence of European Framework Programmes on scientific collaboration in nanotechnology. *Scientometrics* **2013**, *97*, 59–74. [CrossRef]
17. Garfield, E. Journal impact factor: A brief review. *Can. Med. Assoc. J.* **1999**, *16*, 1979–1980.
18. Saha, S.; Saint, S.; Christakis, D. Impact factor: A valid measure of journal quality? *J. Med. Libr. Assoc.* **2003**, *91*, 42–46. [PubMed]
19. Misteli, T. Eliminating the impact of the impact factor. *J. Cell Biol.* **2013**, *201*, 651–652. [CrossRef] [PubMed]
20. Oliveira, E.A.; Peicots-Filho, R.; Martelli, D.R.; Quirino, I.G.; Oliveira, M.C.L.; Duarte, M.G.; Martelli-Junior, H. Is there a correlation between journal impact factor and researchers' performance? A study comprising the fields of clinical nephrology and neurosciences. *Scientometrics* **2013**, *97*, 149–160. [CrossRef]
21. Finegold, L. Writing for science as scholarly communication. *J. Sci. Educ. Technol.* **2002**, *11*, 255–260. [CrossRef]
22. Alaimo, P.J.; Bean, J.C.; Nichols, L. Eliminating lab reports: A rhetorical approach for teaching the scientific paper in sophomore organic chemistry. *WAC J.* **2009**, *20*, 17–32.

23. Robinson, M.S.; Stoller, F.L.; Costanza-Robinson, M.S.; Jones, J.K. *Write Like a Chemist*; Oxford University Press: New York, NY, USA, 2008.
24. Rodriguez, A.C. Teaching peers to talk to peers. *BioEssays* **2012**, *34*, 918–920. [CrossRef] [PubMed]
25. Lebrun, J.-L. *Scientific Writing: A Reader and Writer's Guide*; World Scientific Publishing Co Pte. Ltd.: Singapore, 2013.
26. Derntl, M. Basics of research paper writing and publishing. *Int. J. Technol. Enhanc. Learn.* **2014**, *6*, 105–123. [CrossRef]
27. Conrad, S. Investigating academic texts with corpus-based techniques: An example from biology. *Linguist. Educ.* **1996**, *8*, 299–326. [CrossRef]
28. Assessment Criteria and Level Definitions. Available online: http://www.ref.ac.uk/2014/panels/assessmentcriteriaandleveldefinitions/ (accessed on 29 March 2018).
29. Gunawardena, C. The present perfect in the rhetorical divisions of biology and biochemistry journal articles. *Engl. Specif. Purp.* **1989**, *8*, 265–273. [CrossRef]
30. Stoller, F.; Jones, J.; Costanza-Robinson, M.S.; Robinson, M.S. Demystifying disciplinary writing: A case study in the writing of chemistry. *Across Discip.* **2005**, *2*. Available online: https://wac.colostate.edu/atd/lds/stoller.cfm (accessed on 3 July 2018).
31. Burnard, P.; Gill, P.; Stewart, K.; Treasure, E.; Chadwick, B. Analysing and presenting qualitative data. *Br. Dent. J.* **2008**, *204*, 429–432. [CrossRef] [PubMed]
32. Gill, P.; Stewart, K.; Treasure, E.; Chadwick, B. Methods of data collection in qualitative research: Interviews and focus groups. *Br. Dent. J.* **2008**, *204*, 291–295. [CrossRef] [PubMed]
33. Biber, D.; Conrad, S.; Reppen, R. Corpus-based approaches to issues in applied linguistics. *Appl. Linguist.* **1994**, *15*, 69–189. [CrossRef]
34. Biber, D.; Conrad, S. *Register, Genre and Style*; Cambridge University Press: Cambridge, UK, 2009.
35. Gustavii, B. *How to Write and Illustrate a Scientific Paper*; Cambridge University Press: Cambridge, UK, 2008.
36. Blackwell, J.; Martin, J. *A Scientific Approach to Scientific Writing*; Springer: New York, NY, USA, 2011.
37. Okulicz-Kozaryn, A. Cluttered writing: Adjectives and adverbs in academia. *Scientometrics* **2013**, *96*, 679–681. [CrossRef]
38. Hosgood, G. How to write and publish a scientific paper. *Austral Vet. Pract.* **2011**, *41*, 137–144.
39. Ware, M.; Mabe, M. *STM: An Overview of Scientific and Scholarly Journal Publishing*; International Association of Scientific, Technical and Medical Publishers: Oxford, UK, 2009.
40. White, T. Principles of good research and research proposal guide. In *Policy, Performance and Quality Assurance Unit*; London Council, Borough of Richmond upon Thames: London, UK, 2006; pp. 1–5.
41. Maffioli, F.; Augusti, G. Tuning engineering education into the European higher education orchestra. *Eur. J. Eng. Educ.* **2003**, *28*, 251–273. [CrossRef]
42. Schegloff, E.A. In another context. In *Rethinking Context: Language as an Interactive Phenomenon*; Duranti, A., Goodwin, C., Eds.; Cambridge University Press: New York, NY, USA, 1992; pp. 193–227.