 doi:10.20944/preprints201803.0270.v1

*Article*

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

Myrto-Panagiota Zacharof*[a] and Anna Charalambidou[b]

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

[b] Media Department, Faculty of Arts and Creative Industries Middlesex University London, The Burroughs, Hendon, London, NW4 4BT, UK

*****Corresponding author : [a] Centre for Cytochrome P450 Biodiversity, Institute of Life Science, Swansea University Medical School, Swansea, Wales SA2 8PP, UK myrtozacharof1981@yahoo.com, M.Zacharof@swansea.ac.uk Tel: +44(0) 7455858883

[b]Media Department, Faculty of Arts and Creative Industries, Middlesex University London, The Burroughs, Hendon, London, NW4 4BT, UK , a.charalambidou@mdx.ac.uk Tel: +44(0) 7904365654

**Abstract:** The combination of increased pressures for high-volume, high-impact publications in English language with the high rejection rates of submitted manuscripts for publications presents an often unsurpassable obstacle for (early career) researchers. At the same, the register requirements 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 linguistic choices in 60 published manuscripts in four journals, with impact factor (IF) above 2; all 4 journals, publish original research papers in the field of chemical engineering science and specifically focus on wastewater treatment. Our survey shows that chemical engineering research publications tend to comply to a set of unwritten requirements: multidisciplinarity, brevity, co-authorship, focus on the description of practical results (rather than methods), and awareness of non-specialised audiences. It is found that less discipline-specific vocabulary was used in higher IF journals and this is interpreted within the current context of manuscript publication and consumption. Also, a complex relationship between the advertised scope of each journal and the actual published papers exists, indicating that guide for authors and aims and objective published by the journal's editorial office should be critically evaluated.

**Keywords:** chemical engineering; journal publications; lexical choices; collocations; impact factor; training

## 1. Introduction

Chemical engineering science is a versatile, multifaceted scientific field integrating physics, mathematics, biology and chemistry. Chemical engineers employed both in the academic world and in industry are called to act upon a wide variety of subjects, 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 information, process and evaluate them, take practical steps and make comments and finally communicate these findings in a concise form [2,3]. The prevailing form of communication of research -and therefore its certification- is scientific journal publication, while 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, on the whole, 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 the manuscripts' writing process and its peer review. Moreover, high rejection rates of submitted manuscripts for publications have been observed, with 62% of published paper 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), limited time between submission and publication [9], but mainly 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, employment and funding bodies [39,40]. 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 great number of predatory publishers [10,11], 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 contributed to the widening importance of assessing the value of a publication by (a) the quality of the journal described by the journal's impact factor and (b) the individual citations the publication 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 has been originally designed as an aid to librarians all over the world, to select journals that were most relevant to the public the library addresses or aims to address [13, 14]. It is a ratio calculated by the total number of citations a journal receives over the preceding two years divided by the total number of citations of articles published during that time [15,16,38]. Nowadays, impact factors have been converted to a vital part in decision making regarding scientific impact [14] influencing decisions regarding career prospects, recruitment and appointments [1,16, 17]. Therefore, mastering the art of scientific writing is of utmost importance for every researcher [5] since, research scientists are requested to produce publications of exceptionally high standards, not only related to the novelty and validity of the results presented, but also in a style that would make the manuscript a good read, hence enhancing the potential of publication (by reducing editor's time) and increasing its citability potential [18].

Despite its importance and even though some writing-related training across the curriculum at student level exists [4,19,20], training scientists in publishing their research findings is not an elemental part of chemical engineering education [18]. Post-doctoral researchers are expected to have already obtained the skills required for formulating high quality publications, presentations or talks during their postgraduate education or to have learnt by osmosis, ergo reading published manuscripts from fellow researchers, a tactic that might be highly ineffective, time consuming and lead to failed attempts to publish [21]. Several books have been published offering guidelines for writing papers [22] in science, chemistry and engineering however these give general advice on the structure the papers need to have related to the analysis of experimental data of quantative and qualitative nature without focusing on the use of language [23]. 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)

 doi:10.20944/preprints201803.0270.v1

[20,24-26]. However, the issue of content and register variation among articles published at 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 linguistic choices in 60 published manuscripts in four different journals of impact factor above 2. 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 (Fig. 1). This study explores possible links (or lack thereof) between the impact factor and scope of each journal on the one hand and 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 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, career prospects and attract citations and possible collaborations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

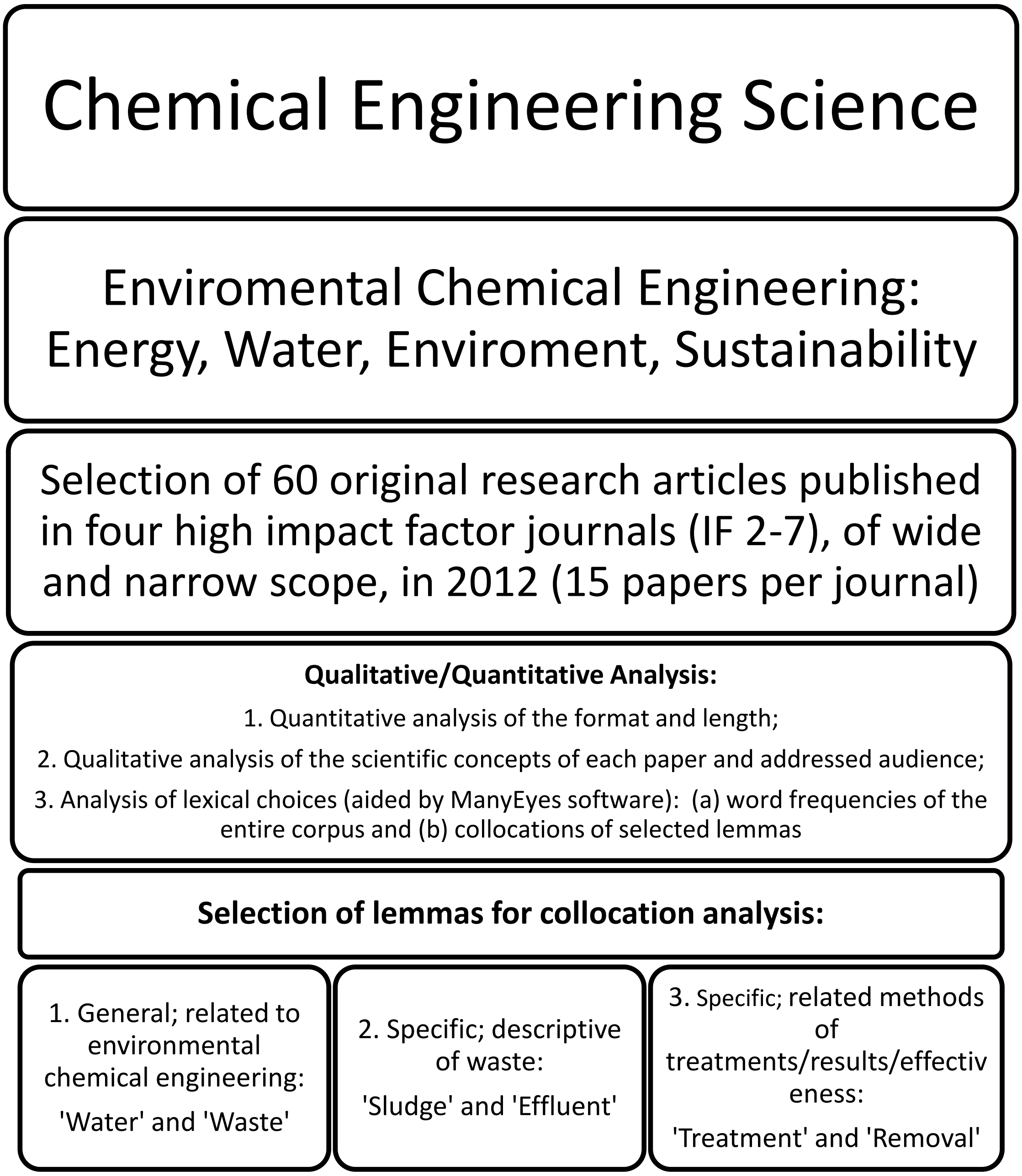



Fig. 1: Schematic representation of the methodology developed and followed in this case study.

## 2. Materials and Methods

### 2.1.Materials

Four journals related to chemical engineering, with impact factor above two have been 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 in 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. Journal's impact factor above 2, considered 2- and 3star, the quality is recognised internationally in terms of originality, significance and rigour [42]

**CES** and **CEJ** were considered journals of wider scope; due to the great variety of scientific categories within chemical engineering from which they accommodate publications (Table 1), while **WR** and **WM** were regarded as specialised scope due to their more concentrated focus on areas relevant to environmental chemical engineering. Each journal publishes various types of papers related to environmental chemical engineering and its major areas of energy, water, environmental impact and sustainability (Table 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

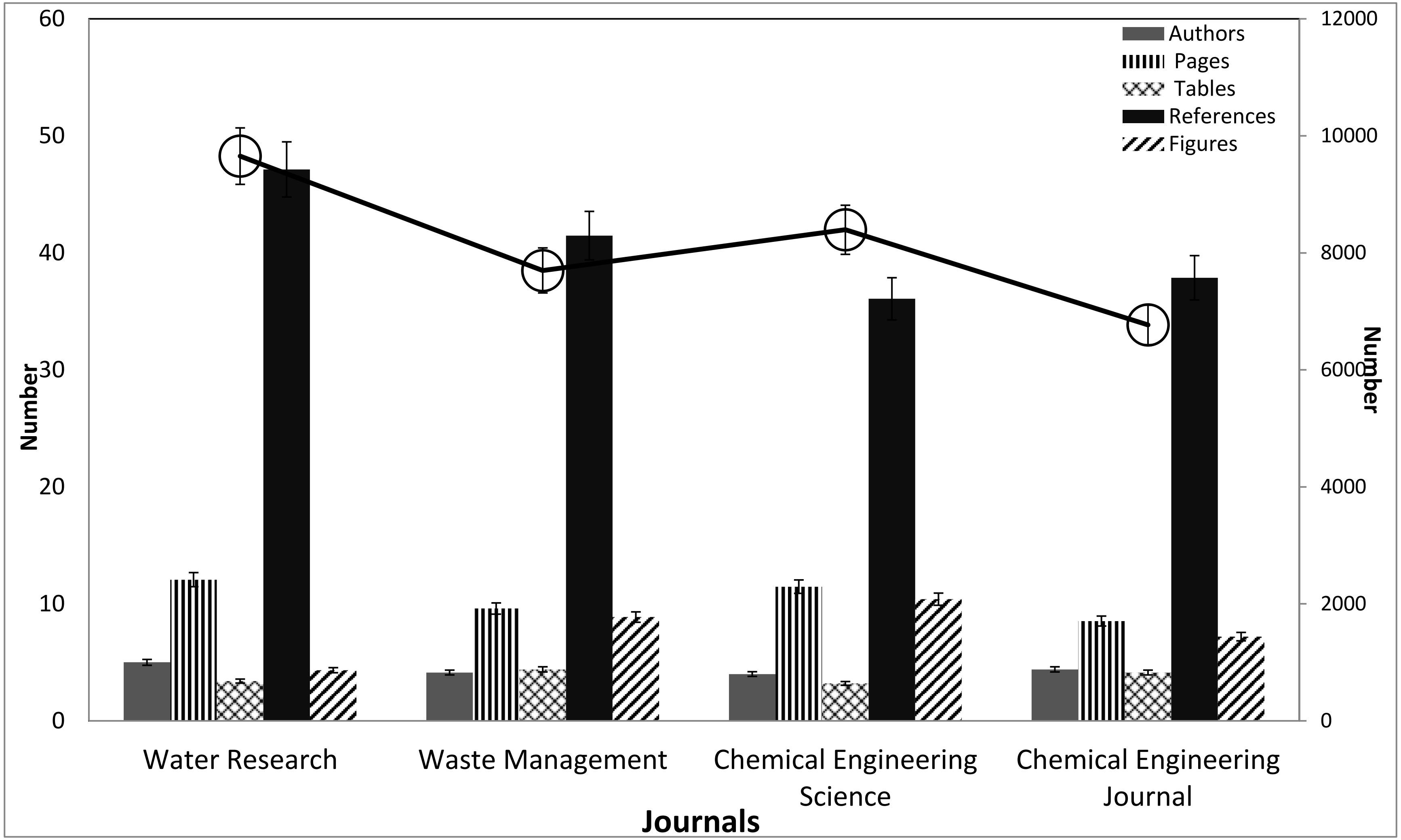



Fig. 2: Volumetric characteristics of the analysed published papers.

15 original papers, i.e. research-related scientific manuscripts describing, analysing and discussing experimental trials and case studies were selected, of every journal totalling in 60 papers, all of them published in 2012. We chose to focus on volumes published 2012, as that would give us a period of five years to track accumulate citations. 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.

### 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 (Fig. 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 freepdf solutions (www.freepdfsolutions.com). Further analysis was done using 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 [27,28] 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 among the reviewers, 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). This software allows for the creation of visualisation 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):

- Tag clouds: visualizations 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.

Many Eyes software can account for large amounts of text and provide accurate and fast calculations, reducing researcher's bias. It can highlight the contrast between our intuitions

 doi:10.20944/preprints201803.0270.v1

about word use and actual patterns in authentic language. An additional benefit is that it has the potential of finding exceptional cases. For the analysis and interpretation of word frequencies and collocations in the various journals analytical tools from corpus linguistics were employed [29,30].

# 3. Results

In order to better contextualize 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 targeted audience.

## 3.1. Format of the collected papers

In practical sciences, including chemistry, physics or engineering, manuscripts are generally considered shorter in length compared to liberal sciences and arts [31]. Commonly within a breadth of 6 to 12 printed two-column pages, including tables, figures and references [32] the authors are expected to satisfactory demonstrate and explain their reason for research and findings. Reduction in volume and size of research papers have been implemented unanimously to physical sciences journals due to the constantly increasing rate of submission, leading to the need to accommodate a higher number of published papers within journals printed issues or volumes [33]. Shorter length of such papers is also supported by the ability of the authors to visualize their findings into meaningful figures that need little or no explanation as well as reducing the amount of words and development of long, articulate arguments by tabulating their core finding [34].

These findings are also supported in this case study. The papers’ length was between 8 to 13 printed pages, including figures and tables, with a word count between 6800 to 9700 words including references, highlights, abstracts and tables and figures legends (Fig. 2). As regards to the length of the papers, similarities were found between **CES** and **WR** (average 11 pages, 8.300 words) and **CEJ** and **WM** (average 8 pages, 6.700 words). Cited literature serves in supporting the findings and explaining the reasoning behind the trials, but also saving space, as the authors are not forced to refer extensively to previously developed knowledge. References in all papers ranged between 36 to 48, with similar trends found among the wider scope journals **CES** and **CEJ** (on average 37 references) and the specialized scope **WM** and **WR** (on average 41 references) (Fig. 2).

Figures and tables are the core part of the published manuscripts, varying in numbers, 4 to 10 figures and 3 to 5 tables, proving essential for the understanding and scientific evaluation of the papers. Within that context, the text serves for analyzing, explaining and discussing these visual aids to the audience. Papers in **CEJ** and **WM** were small in size, quite densely written, and comprising mainly graphs and figures without analytically describing numerical results. **CES** and **WR** publish longer papers with numerous figures and analytical numerical data, encouraging elaboration and explanation of findings while WR has a balance between figures, tables and discursive sections.

The quantitative analysis suggests that the selected papers from each of the four journals share similar quantitative characteristics, thus rendering the four datasets comparable.

### 3.2.Multidisciplinary nature of the analyzed 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 to prove the scientific concept and reasoning developed in the manuscripts (Ware and Mabe, 2009; White, 2006). The multifaceted nature of the published papers in chemical engineering was clearly reflected in this study, by the subject category (Fig. 3) and audience distribution (Fig. 4).

Out of the 60 papers investigated, the array of subjects of interests relevant to environmental chemical engineering and specifically to waste treatment and management is wide (Fig. 3), covering numerous scientific areas from biochemical engineering to environmental chemistry, to other engineering disciplines such as mechanical, electrical or civil engineering. The two most often-encountered areas were environmental engineering (up to 35%) and wastewater treatment (up to 30%) making these two (Fig. 3), while a more general approach to biochemical and chemical engineering related paper was the next prevalent subject area (up to 23%). Solid waste treatment (up to 26.7%) and chemistry (up to 20.93%) are also covered in the journals. When compared to the advertised scientific subject of interest for publication of each journal, a differentiation is found since the advertised subject areas are broader to the categories that emerged from this research. 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 lexical visualisation can provide a quick way for researchers to assess the specific areas that are most likely to be published in the journal.

 doi:10.20944/preprints201803.0270.v1

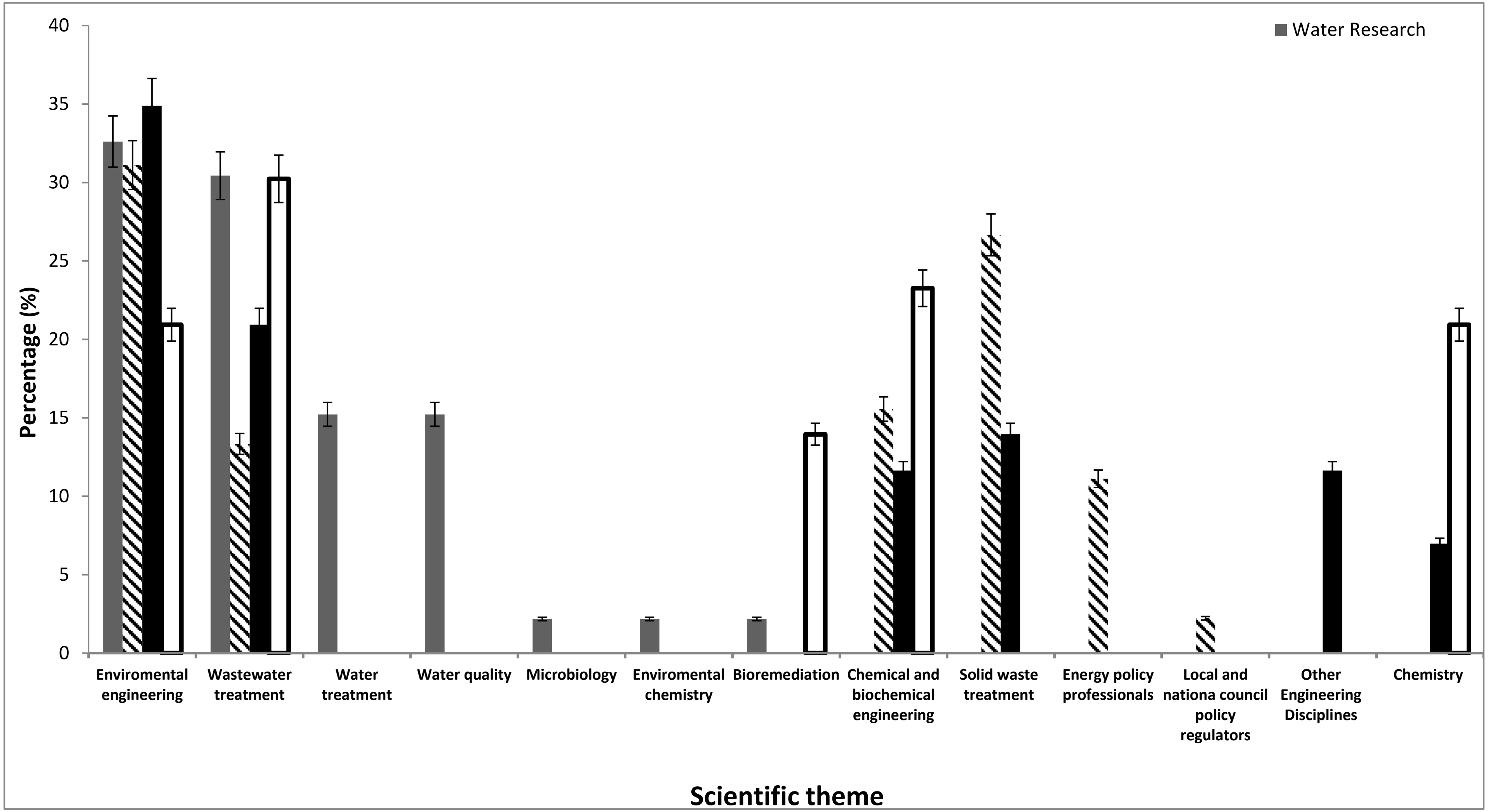



Fig.3. Scientific theme distribution of the analysed published manuscripts.

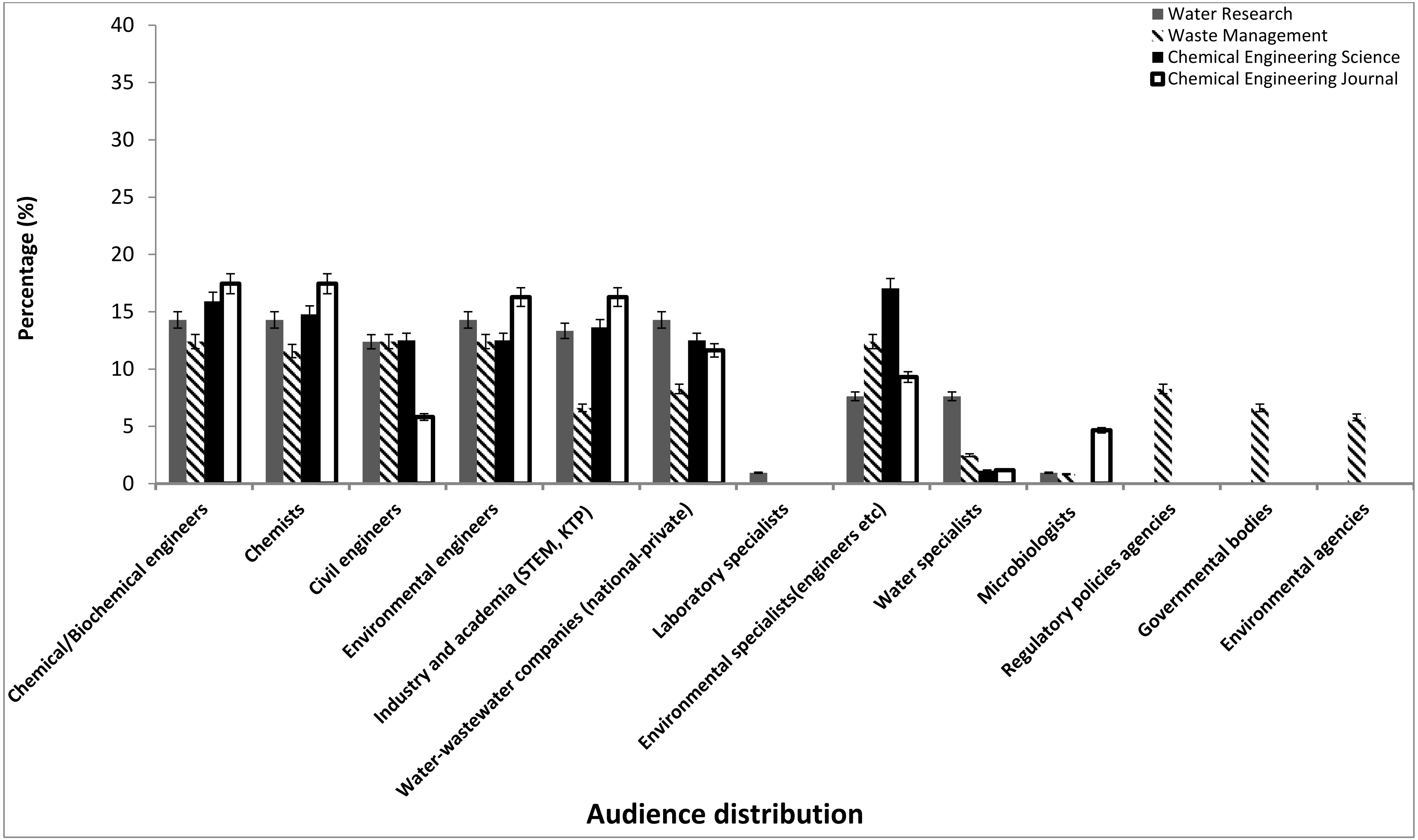



Fig. 4: Distribution of potential audience of the analysed published manuscripts

The topics covered in the papers in our sample was 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 13 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 (Fig. 4). In particular, 8 of these categories are represented in all the selected journals into varying percentages (5.81% to 17.5%). This is a divergence from the advertised audience in the website of each journal, where the focus is on specialist in chemical engineering audience within the field.

The content analysis of the papers has shown that in **CEJ** and **CES** there is a stronger tendency, compared to **WR** and **WM**, to appeal to the industry. That could be attributed to the nature of studies, i.e. dealing with trials in pilot plant scale (large volumes of materials), which are more attractive to the industry, since the authors have not only proven their concept but have also implement 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; since a more holistic approach is taken that accounts for financial and social parameters. Thus patterns have emerged about the nuances of the addressed audience in the published manuscripts of each journal, which are not clearly communicated in the journals' websites.

The wide range of potential audience 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. Figures and tables might, to a certain extent, describe the core essence of the paper but the text, especially in the discussion and conclusions part are vital for the overall understanding of the ideas. This is found also in this case study, where the words "table" and "figures" are among the top ten words mostly used 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.

### 3.3. Analysis of lexical choices

As suggested in the introduction, linguistic and in particular 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 collocations of certain key words, and correlations were explored between the results and the type of journal (wider or specialized scope) and the journal's IF. Six lemmas were chosen, to explore collocations and consequently the context in which certain key terms are employed and variation in the specific meaning that is ascribed to them (Fig. 1). These terms comprise ***water*** and ***waste,*** which are generally used when referring to the environment and would be expectably mentioned mostly in the introduction and discussion or conclusions parts of the papers, two lemmas specifically related and descriptive of waste, ***sludge*** and ***effluent***, that 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 Fig. 5, for the frequency of occurrence of these six lemmas in each journal). The collocations 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 [41]) were analysed based on ‘word trees’ and ‘phrase net’ visualizations (see supplementary material). Below the key findings of the analysis of the ManyEyes 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 visualization 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 sterilization, as solvent, as treatment method for other elements or as a component to other substances.

As regards to the 15 **WR** and 15 **WM** papers, “water”, was found 792 and 244 times respectively, and, as its collocations suggest, it was conceptualized 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”, found 413 times, in **CEJ** 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.

Similarly to the use of the lemma “water”, “waste”, occurring 131 times, in **CES**, was mainly conceptualized as part or a tool of 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, similarly 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.

Both lemmas “water” and “waste” were routinely found in the manuscripts of each journal, and they were among the top 10 words most often-encountered words in the manuscripts, and used in high frequency either combined, i.e. wastewater, or separately (Table 2). However, as the analysis above indicates, in **CEJ** and **CES** the terms were recurrently employed in different contexts than **WR** and **WM**. In **CEJ** and **CES** the words are used in a rather specialized 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

anticipated that a less restrictive use of the term would have been encountered. In **CEJ** and **WR**, the words are found in analogous amounts; while in **CES** the amount of use is very limited, suggesting the use of a scientific specialized vocabulary (e.g. the terms “liquid” or “fluid” or “solvent”, were preferred over “water”). On the other hand, **WM** is standing 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 use of the words “sludge”, “effluent”, “treatment” and “removal” (Fig. 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 of waste. On the other hand, in **CES** “sludge”, found 205 times, was used to describe a muddy, murky, highly viscous thick material in the need of processing or treatment 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 as murky, muddy, soil based material, liquor or concentrated liquid of a semisolid nature coming out of process treating sludge.

In **CEJ** and **CES** the word “effluent” is not found, implying the absence of 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 relates 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

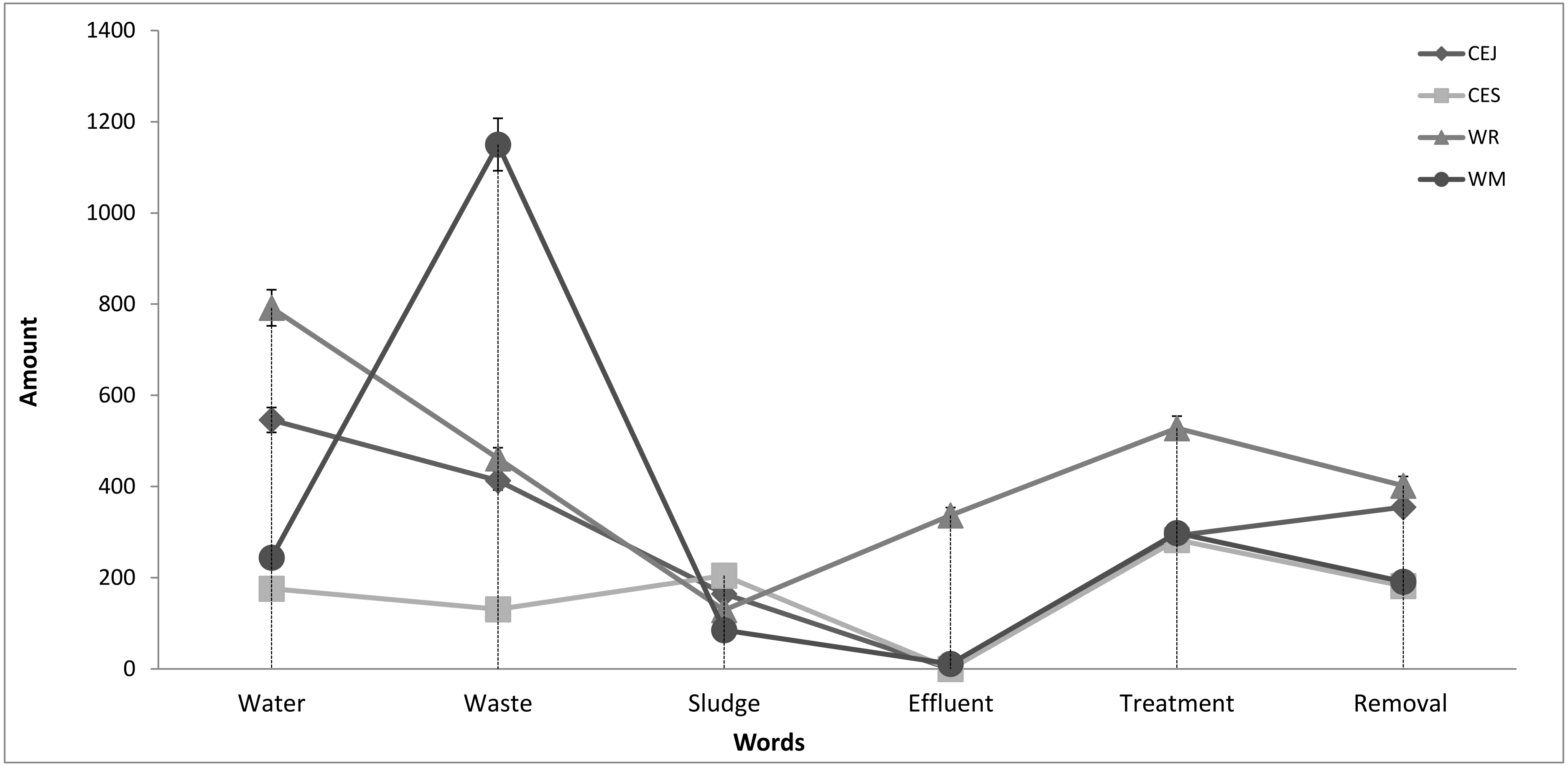
 doi:10.20944/preprints201803.0270.v1

explain any liquid coming out, discharged of a waste treatment or of waste producing process, while in **WM** it occurred only 11 times, and was used when referring to 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 correspondingly, 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, and its effectiveness and efficiency. In **CES** "removal" also represented the main scope of the project developed in the manuscript. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

131 


132 Fig. 5: Distribution of the selected keywords among the analysed manuscripts

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

## 4. Discussion

The analysis of visualisations and word frequencies (see supplementary material) has shown emerging trends in lexical choices that also have implications about the specific subject-area and approach preferred in each journal and which -interestingly- do not necessarily correspond with the advertised scope of each journal or with the authors' expectations. Among the four journals selected, **WR** and **WM** were considered of specialized scope (based on their advertised scope), thus expected to accommodate a highly specialized 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 a more specialized vocabulary is used, especially in **WM**. The high occurrence of discipline-specific vocabulary is not only associated with the scale of the experiments, but also with 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 indicates (Table 2). This finding in also supported by the close reading of the published manuscripts, that has revealed that the manuscripts in **WM** and **CES** refer to specialized and complex methods of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

chemical engineering (Fig. 3 and Fig. 4). For example in **CES** instead of plainly using "water" other related terms are used such as "concentrations", "phase" which point towards to chemical processing, whereas in **WM** terms related to water such as "leachate" are used to point residuals of solid wastes.

| **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** | 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** | 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** | Three aspects of chemical engineering: **chemical reaction engineering, environmental chemical engineering, and materials synthesis and processing.** |

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

| | | | | |
|---|---|---|---|---|
| | **analysis by chemical, physical and biological methods** | **considerations, financial and marketing aspects, policy and regulations, education and training, planning and implementation.** | **systems engineering reactions, separations science and technology** | |
| **Abstract** | **Concise and factual, descriptive (up to 250 words)** | | | |
| **Graphical abstract** | **Optional** | | | |
| **Highlights[2]** | **Mandatory** | | | |

**Table 1**: Summary description the prerequisites request by each journal for the submission of manuscripts based on of the full aims and scope and guide for authors, publically available on the journals' websites.

[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.

On the other hand, CEJ and WR use a less discipline-specific vocabulary, with salience of terms “wastewater”, “effluents”, “samples” which are far less frequent in **CES** and **WM** (Table 2). **WR** published papers are indicating a holistic approach to water-related research focusing on the findings of the experimental trials and their applicability in the society, addressing social, financial and legal aspects. This also corresponds with the frequent use of the lemmas “removal” and “environmental”.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

| **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 | **wastewate**r | 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)** | 144798 | 115491 | | 101519 | | 125910 | |

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

In the case of **CEJ** the findings, from the qualitative analysis of the papers, regarding the multidisciplinary nature of the papers are mirrored in the results of lexical analysis. It further confirms that the use of a less discipline-specific vocabulary enhances the readability of the journal, which can reach a wider audience, including industrials and policy regulators. On the other hand, while **WR** has been found in the qualitative analysis to target in the main an academic audience, the use of simplified vocabulary boosts its readability among scientists from a wide range of varying disciplines.

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 of 4.655 and 3.473 respectively. Technology has facilitated tremendously knowledge exchange shifting from only printed media to a combination of available online, easily downloadable articles and printed media, expanding significantly the availability of a paper, as the readers are not depended only on the printed resources that exist in libraries and repositories across the world [17, 37].

Literature searches are not necessarily guided by advisors, supervisors or assisting librarians, and are being partially replaced by specialized 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 reading of the majority of published papers, on an individual unsupervised basis, from an audience that may not have an extensive knowledge on the subject (postgraduate students, early career, professionals, researchers, academics and fellows), and 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 can 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 are found. Among all four journals, only **WR** published papers reflect the journal's very broad approach, focusing on innovation without disregarding new approaches to current techniques. **CEJ** and **CES** have a narrowed thematology, addressing highly specific subjects contrary to the journals advertised spectrum. In the published manuscripts, emphasis is placed on optimization of existing methods, mainly chemical treatments rather than innovation, which cannot be as easily and quickly applicable. A similar tendency is found in **WM**, where, in spite of the advertised wide array of publishing subjects, the published manuscripts do not cover such a wide spectrum, and focus 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 generalizable across journals of practical sciences, let alone all disciplines. However, due to the depth of the investigation this snapshot of trends in published chemical engineering research has offered an insight on 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

learning and teaching material for chemical engineers, but also researchers from other disciplines that seek to publish their research include the following.

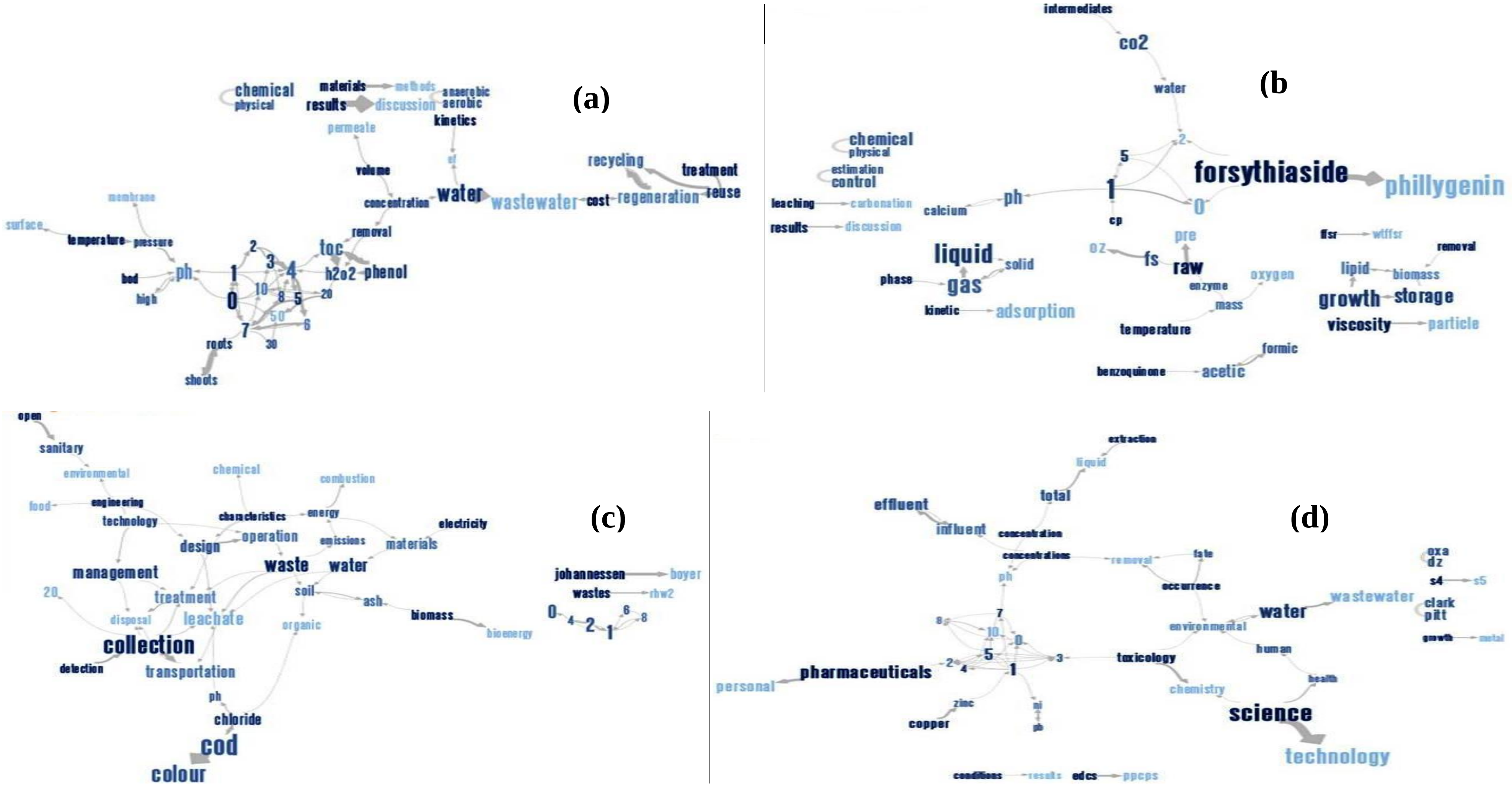
- Highly discipline-specific vocabulary use, including extensive use of acronyms, should be avoided where possible, to aid favorable consideration of manuscripts at higher IF journals and to increase the citability potential of the article.
- There is a complex relationship between the thematology, the audience and the scope, as they are advertised in the journal's website, and the actual published manuscripts.
- Guide for authors and journal aims and objectives, published by the journal's editorial office, should be taken into account, to help authors make an initial decision regarding the journal that is most suitable for the submission of their research, but should be critically viewed.
- It is recommended for prospective authors to collect a number of publications, of their journal of choice, published within close proximity, to the potential submission date, in order to get a better understanding of the journal's thematology, the approaches favored and preferred discourse style.
- Visualisations of word choices and associations, which can be fairly easily and quickly done with the aid of freely available software, is a very powerful tool in providing an accurate overview of both the preferred content andapproach of each journal, as well as its preference as regards to lexical choices. They can be an indispensable tool for chemical engineering students and novice researchers that wish to gain an emit understanding of the actual scope of the plethora of journals within each discipline, without having to engage in the labor-intensive close reading of a large corpus of published papers.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

Extending this research to similar investigations of a larger size of text samples, representing more fields of science would be desirable, so that the findings will then be more representative of scientific writing in English. Further exploration of links between linguistic 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 the researchers to write in a style that best suits their target journal.

 doi:10.20944/preprints201803.0270.v1

**Supplementary Materials:**



*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

Fig.1: Phrase nets graphical images depicting the collocations between the selected words for analysis and the remaining words in the selected published manuscripts in (a)CEJ, (b)CES, (c)WR, and (d)WM provided by the lexical visualisation software Many Eyes and used for the qualitative analysis of the published manuscripts in this case study.

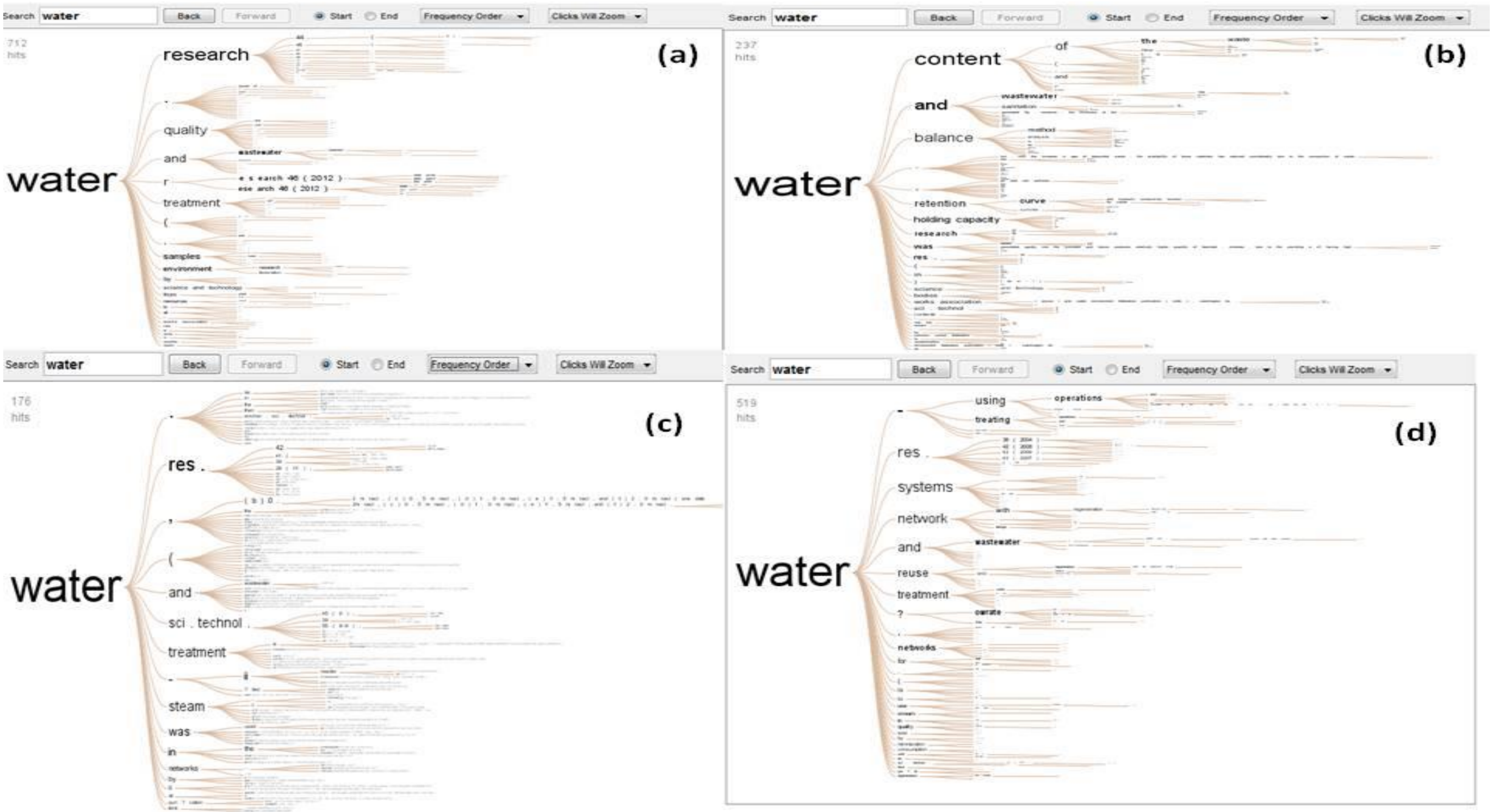



Fig.2: Word trees depicting the word "water" in (a)CEJ, (b)CES, (c)WR, and (d)WM and its collocations (word associations) provided by the lexical visualisation software Many Eyes and used for the qualitative analysis of the contexts of use of selected lemmas in the published manuscripts in this case study.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 March 2018 doi:10.20944/preprints201803.0270.v1

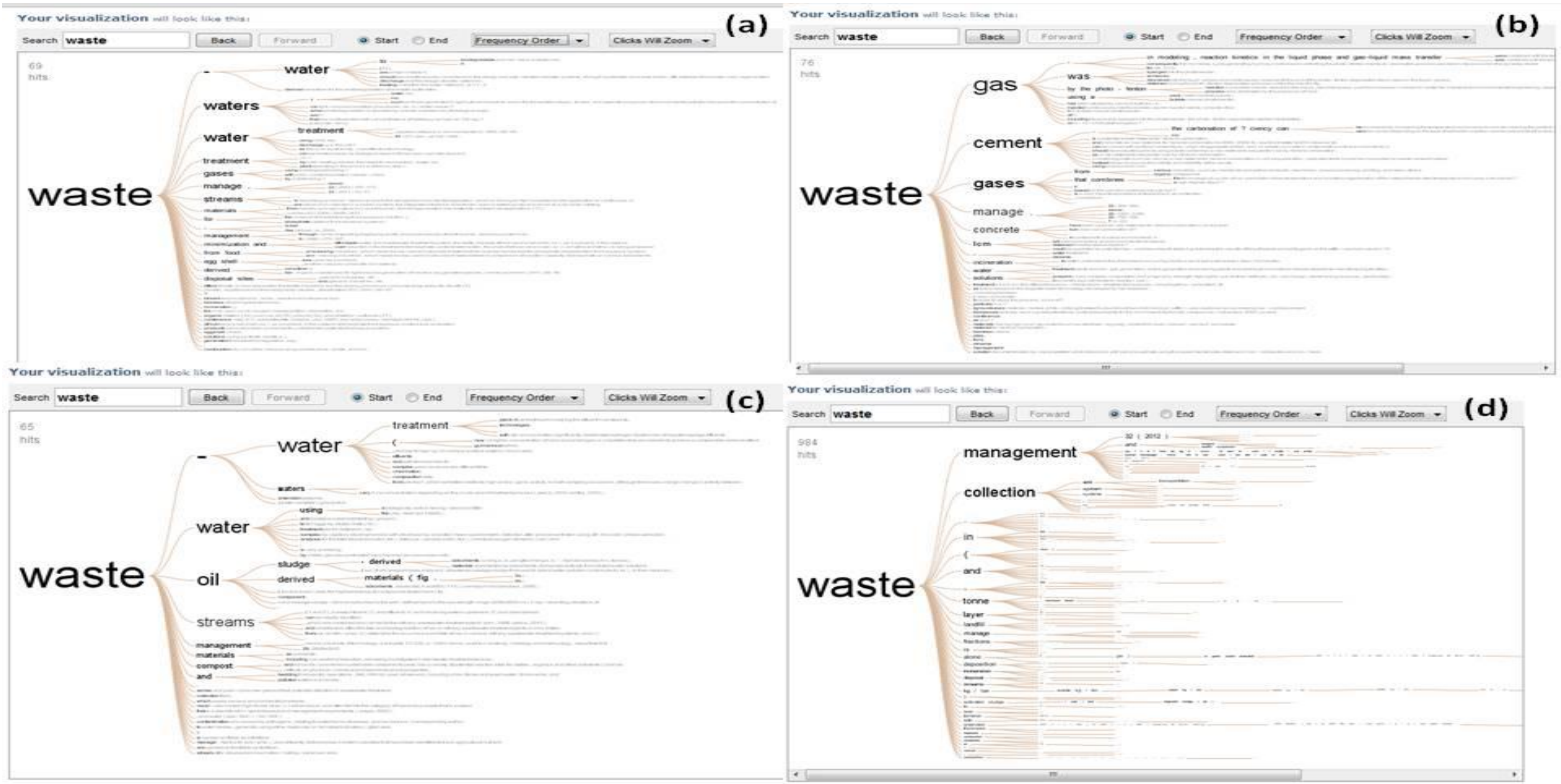



Fig.3: Word trees depicting the word "waste" in (a)CEJ, (b)CES, (c)WR, and (d)WM and its collocations (word associations) provided by the lexical visualisation software Many Eyes and used for the qualitative analysis of the contexts of use of selected lemmas in the published manuscripts in this case study.

**Acknowledgments:** The authors would like to thank fellow researchers in the College of Engineering, Swansea University for their valuable advice in the research discussed here.

## References

[1] Abramo, G., D' Angelo C.A., Rosati F., 2014. Career advancement and scientific performance in universities. Scientometrics, 98, 891-907.

[2] Massoudi, M., 2003. Can scientific writing be creative? J. of Sci. Edu. Tech., 12, 115-128.

[3] Vanclay, J.,2012. Impact factor: outdated artefact or stepping-stone to journal certification? Scientometrics. 92, 211-238

[4] Beall, H., 1998. Expanding the scope of writing in chemical education. J. Sci. Edu.Tech., 7, 259-270.

[5] Uskokovic, V. 2012. Ten Commandments for writing a meritable scientific paper. J. Postdoc. Affairs, 2, 2-7.

[6] Bosquet, C., Combes, P.-P., 2013. Are academics who publish more also more cited? Individual determinants of publication and citation records. Scientometrics. 97, 831-857

[7] Porter, A., Rafols, I., 2009. Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics. 81, 719-745.

[8] Woolley, K., Barron, P., 2009. Handling manuscript rejection insights from evidence and experience. Chest. 135,573-5

[9] Drummond, C. W. E., Reeves, D., 2005. Reduced time to publication and increased rejection rate. J. Antimicrob. Chemo. 55, 815-816.

[10] Beall, J., 2012. Beall's list of predatory publishers 2013, *http://scholarlyoa.com/publishers* (last accessed 29 Mar 2018).

[11] Moore, A. (2009). The garbage collectors - Could a particular sector of author-pays journals become silently acknowledged collectors of scientific waste? *BioEssays*, *31*, 821.

[12] Amin, M., Mabe, M., 2000. Impact factor: use and abuse. Persp.s in Publ., 1, 1-6.

[13] Fassoulaki, A., Sarantopoulos, C., Papilas, K., Patris, K., Melemeni, A., 2001. Academic anaesthesiologists' views on the importance of the impact factor of scientific journals: a North American and European survey. Gen. Anaesthesia. 48, 953-957.

[14] Ovalle-Peradones, M.-A., Gorraiz, J., Wieland, M., Gumpenberger, C., Olmeda-Gomez, C., 2013. The influence of European Framework Programmes on scientific collaboration in nanotechnology. Scientometrics 97, 59-74.

[15] Garfield, E., 1999. Journal impact factor: a brief review. Can. Med. Assoc. J. 16, 1979-980.

[16] Misteli, T.,2013. Eliminating the impact of the impact factor. J. Cell Biol. 201, 651-652.

[17] Oliveira, E. A., Peicots-Filho, R., Martelli, D. R., Quirino, I. G., Oliveira, M. C. L., Duarte, M. G., Martelli-Junior, H., 2013. Is there a correlation between journal impact factor and researchers' performance? A study comprising the fields of clinical nephrology and neurosciences. Scientometrics. 97, 149-160.

[18] Finegold, L.,2002. Writing for science as scholarly communication. J. Sci. Edu.Tech. 11, 255-260.

[19] Alaimo, P.J., Bean, J.C., Nichols, L., 2009. Eliminating lab reports: a rhetorical approach for teaching the scientific paper in sophomore organic chemistry. WAC J., 20, 17-32.

[20] Robinson, M.S., Stoller, F. L.,Costanza-Robinson, M.S., Jones, J.K., 2008. Write like a chemist. New York: Oxford University Press.

[21] Rodriguez, A.C., 2012. Teaching peers to talk to peers. BioEssays. 34, 918-920.

[22] Lebrun, J.-L., 2013. Scientific writing: a reader and writer's guide. Singapore: World Scientific Publishing Co Pte. Ltd.

[23] Derntl, M., 2014. Basics of research paper writing and publishing. Int. J. of Tech. Enh. Learn 6, 105-.123

[24] Conrad, S., 1996. Investigating academic texts with corpus-based techniques: an example from biology. Ling Edu. 8, 299-326.

[25] Gunawardena, C., 1989. The present perfect in the rhetorical divisions of biology and biochemistry journal articles. Engl. Spec. Purp., 8, 265-273.

[26] Stoller, F., Jones, J., Costanza-Robinson, M. S.,Robinson, M. S., (2005). Demystifying disciplinary writing: A case study in the writing of chemistry. *Across the Disciplines, 2*. http://wac.colostate.edu/atd/lds/stoller.cfm (last accessed 15 April 2015).

[27] Burnard, P., Gill, P., Stewart, K., Treasure, E., Chadwick, B., 2008. Analysing and presenting qualitative data. Brit. Dent. J. 204, 429-432.

[28] Gill, P., Stewart, K., Treasure, E., Chadwick, B., 2008. Methods of data collection in qualitative research: interviews and focus groups. Brit. Dent. J. 204, 291-295.

[29] Biber, D., Conrad, S., Reppen, R., 1994. Corpus-based approaches to issues in applied linguistics. App. Ling., 15, 69-189.

[30] Biber, D., Conrad, S., 2009. Register, genre and style. Cambridge, UK; New York: Cambridge University Press.

[31] Gustavii, B., 2008. How to write and illustrate a scientific paper. Cambridge, UK; New York, Cambridge: Cambridge University Press.

[32] Blackwell, J., Martin, J., 2011. A scientific approach to scientific writing. New York: Springer.

[33] Okulicz-Kozaryn, A., 2013. Cluttered writing: adjectives and adverbs in academia. Scientometrics. 96, 679-681.

[34] Hosgood, G.,2011. How to write and publish a scientific paper. Austral. Vet. Pract., 41, 137-144.

[35] Ware, M.,Mabe, M., 2009. *STM: An overview of scientific and scholarly journal publishing*. Oxford: International Association of Scientific, Technical and Medical Publishers.

[36] White, T.,2006. Principles of good research and research proposal guide. In: Policy, performance and quality assurance unit. (pp. 1-5). London: London Council, Borough of Richmond upon Thames.

[37] Maffioli, F., Augusti, G.,2003. Tuning engineering education into the European higher education orchestra. Eur. J. Eng. Edu.28, 251-273.

[38] Saha, S., Saint, S., Christakis, D.,2003. Impact factor: a valid measure of journal quality? J. Med. Librar. Assoc. 91, 42-46

[39]Barrow, L.H., 2003. Searching for educational technology faculty. J. Sci. Edu.Tech., 12, 143-147.

[40] Tscharntke, T., Hochberg, M.E., Rand, T.A., Resh, V.H., Krauss, J., (2007. Author sequence and credit for contributions in multiauthored publications. PLOS Biol. *5*, 13-14.

[41] Schegloff, E. A., 1992. In another context. In A. Duranti and C. Goodwin (Eds.), Rethinking context: language as an interactive phenomenon. (pp. 193-227). Cambridge; New York: Cambridge University Press.

[42] http://www.ref.ac.uk/2014/panels/assessmentcriteriaandleveldefinitions/ (last accessed 29.Mar.18)