A short review of Research based Teaching and Learning content in Natural Language Generation

KEYWORDS

NLG, Teaching and Learning, Natural Language Generation, Data-to-Text Generation, Qualitative review

ACM Reference Format:
. 2018. A short review of Research based Teaching and Learning content in Natural Language Generation . In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

With respect to the challenges of Big Data Analytics, the vast majority of efforts are currently focused on the analysis and processing of Big Data in order to glean insights for their transformation into actionable knowledge by stakeholders. However, the effective delivery of these insights for decision-making purposes to the non-expert user remains an open challenge. While visualisation may provide the overall analytic view, it does not provide tailored insights (beyond the data analyst) to the average user. Hence, Natural Language Generation (NLG) can be considered as the last hurdle in the Big Data race[1]. Without NLG, Big Data Analytics is heading towards an inevitable knowledge access bottleneck, whereby data analytics will have produced valuable but inaccessible insights to the non-expert. Current efforts in Natural Language Processing (NLP) continue to focus heavily on Natural Language Understanding (NLU) of text. Despite this, one leading research advisory company, Gartner, claims that by "2019, natural-language generation will be a standard feature of 90% of modern BI and analytics platforms"[2]. Moreover, NLG was cited by Forbes in 2017 as number one Top 10 Hot AI technology[3]. Aside from coverage in technology media, NLG has also enjoyed renewed research interest by machine learning (ML) researchers due to advances in deep learning[15, 19]. Taking into account the factors above, it is inevitable that for newcomers (in particular from industry) to the field, their focus will be on accessing NLG software, more importantly, methodologies quickly and easily. Universities, as higher institutions of education which typically transfer such specialist knowledge through instruction, will be expected to meet this demand. Hence, as a first step, it is essential that one first reviews teaching and learning activities with respect to NLG. This paper presents a qualitative scoping exercise, which analyses academic course content for the purpose of classifying and aligning NLG teaching and learning material with corresponding research concepts in the field. We engage in a manual content analysis of NLG course modules in order to gain *knowledge for understanding*[33]. Specifically, we are seeking to gain an overview of academic practice with respect to the transmission content knowledge for teaching and learning activities in NLG and to what extent it is informed by the research literature in the field [21]. The analysis will be of particular use to:

- Developers in industry who are looking to develop NLG applications for their products and are searching for learning materials.
- Researchers who are new to the field and wish to engage in the scholarship of NLG.
- Academics who have been tasked to design new NLG courses including, learning outcomes, teaching & learning and assessment activities.
- Academics who wish to design a new teaching and learning session on NLG into an existing NLP course.
- Academics who search to engage in pedagogical research with respect to Teaching and Learning for NLG and/or NLP.

The remainder of the paper is structured as follows: Section 2 describes related work, Section 3 describes our methodology, while Section 4 offers an analysis and discussion. Finally, Section **??** concludes the paper.

2 RELATED WORK

With respect to choosing a research methodology, qualitative analysis, specifically content analysis [16] and similarly grounded theory [14] (See Section 3 for more details) are the appropriate methods given the research goals, time and resources, and the nature of data sources (i.e., course Web pages and documents). In addition, the resource constraints also limit the ability to generate data and engage with other sources i.e., structured interviews with module owners or student surveys, focus groups and/or case studies [16]. While transnational curriculum analysis investigating threshold concepts may at a first glance seem relevant, we do not seek to uncover content knowledge which poses difficulty in the teaching and learning of NLG. Rather, we only wish to capture a general conceptual structure of content knowledge used for teaching and learning of NLG. However, since the analysis of content will likely be informed by research in NLG, it will be necessary to access a systematic review of the relevant research literature, in order to link university course content to relevant research themes. In this study, we exploit predefined categories of common research themes in NLG-based on a recent and comprehensive systemic review in the field of NLG [19]. Using predefined categories are quite typical of content analysis. However, in this study, we also apply some

[1] See WHITE PAPER: The Last Mile in Delivering Information from Big Data An NLG Thought Leadership White Paper by Dr. Robert Dale, NLG Chief Strategy Scientist www.arria.com
[2] https://tinyurl.com/yck3rhoc
[3] See https://tinyurl.com/y7xkal2n

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

grounded theory to the course content and seek to align any emergent categories with respect to subject matter content with research theme content.

Conceptual structures or classification trees have often been applied to capture content knowledge of discipline subject matter [10, 30, 31]. With respect to using conceptual structures in computer science, the obvious parallel was the usage of Testable Reusable Units of Cognition (TRUCs), whereby pedagogical patterns are inspired by software design patterns [24, 27]. There have also been attempts to define the body of knowledge in Computer Science formally using an ontology within the Special Interest Group on Computer Science Education within the ACM (Association for Computing Machinery)[13, 17]. In addition, the ACM Computing Classification, lists NLG as a sub-concept of NLP[4]. Similarly, the ACM Curriculum model, the international scientific and professional society, and the Association for Computational Linguistics[5] defines the core body of Knowledge for Computational Linguistics' curricula which includes NLG and the sub-concepts of text planning and linguistic realisation. However, similar to the ACM, it is not meant as a complete curriculum, but rather a "minimal core consisting of units for which there is broad consensus"[6]. This can also be supplemented with additional material. The Computation and Language (CL) core curricula was a consensus output derived from [11], the last of three workshops ending in 2008 on teaching and learning activities in Computational Linguistics and NLP [1, 3, 5]. To our knowledge, there have been since then no other dedicated research venues with a teaching and learning focus in NLP and furthermore, none of the workshop proceedings contain and specific focus on NLG.

3 METHODOLOGY

The rationale of this study is to understand academic practice with respect to content knowledge for teaching and learning activities in NLG, and to what extent it is informed by the research literature in the field. This study is qualitative and will involve content analysis, whereby the module content under investigation will be analysed, coded and categorised [16]. Observations and patterns will be captured and summarised. The output will be involve the production of a hierarchical conceptual structure or a classification tree (categories and subcategories of content knowledge), which capture the structure of subject matter/content knowledge across each module analysed. Moreover, in the spirit of grounded theory, the process will be inductive and emergent, so a theory may emerge which fits the observed data. The approach is positivist and fits what is exploratory research [14]. The study is effectively a form of documentary analysis[32], whereby we are looking at public online institutional documents, online resources available at the module level i.e., module descriptors, lecturing material, as well as the relevant academic and research publications linking to the institutional evidence. The methodology for this study consists of the following steps:

- **Step 1:** Analyse the most recent systematic review for NLG [19] and engage in content analysis in order to extract a high level conceptual structure of the most up-to-date common research themes in NLG - which we shall call **NLG-CS-Research**.
- **Step 2:** Select the relevant NLG modules (if not all) and gather all content which provides a description of the content knowledge included in the module i.e., module descriptors, lecture slides, reading lists, source code references, assessment and assignment descriptions.
- **Step 3:** Engage in the process of content analysis above, which will be informed by grounded theory as described in [16] i.e., define population, sample, define the context and units of analysis, extract content, sort, cluster, generate topics.
- **Step 4:** Map over concepts from NLG course material to the **NLG-CS-Research** classification tree in order to understand how much current research in NLG is reflected in NLG teaching and Learning.
- **Step 5:** Analyse the results, report observations and offer the conclusion and recommendations.

3.1 Ethical Issues

There were no ethical issues, as the data is publicly available on the Web and nor does it contain any personal (sensitive) data requiring consent from individuals i.e., module owners. No subsequent data protection mechanisms are needed as the data is already public on the Web. Finally, module owners are anonymised, although institutions are listed[7].

3.2 Data Sources

The following data sources are included in the study:

- University approved modules with a focus on NLG at the undergraduate and postgraduate levels.
- Research-led training courses i.e., Summer Schools, Doctoral Summer Schools.

The following data sources are **not included** in this study

- Dedicated tutorials in established conferences within the field.
- NLG Text Books.
- Undergraduate/postgraduate modules in Computer Science that provide only one teaching session in NLG.
- Non peer reviewed sources, such as user generated content from blogs, discussion boards, Quora[8], etc.

Finally, this is not an analysis of the state-of-the-art of NLG research, although the process of reviewing teaching and learning in NLG may yield, as a side effect, themes and trends in NLG research. A series of manual Web searches were carried out against Google Scholar [9], ACL Anthology [10], as well as the email archives of the Special Interest Group on Natural Language Generation (SIGGEN) Mailing List[11].

[4] https://www.acm.org/publications/class-2012
[5] https://aclweb.org/
[6] https://aclweb.org/aclwiki/Core_body_of_knowledge

[7] The above is in compliance with the EU general data protection regulation, which entered into force in May 2018. See Recital 159 at https://gdpr-info.eu/recitals/no-159/, Accessed 16 July 2018
[8] https://www.quora.com/
[9] https://scholar.google.com/
[10] http://aclweb.org/anthology/
[11] https://tinyurl.com/y964n6x3

4 ANALYSIS AND DISCUSSION

4.1 Analysis of NLG University Courses

Table 1 summarises the courses available and their respective data sources. We list **four university** courses which are each dedicated fully to NLG. In addition, we include other shorter specialised courses, which are also university-led. With respect to headings in Table 1, **Course** is a unique identified for reference purposes, while **Discipline** represents the department *owning* the module. In addition, **URL** references the source link on the Web, while **ECTS** indicates the number of credits in accordance with the European Credit Transfer and Accumulation System (ECTS)[12]. Depending on the country, one ECTS credit point can equal on average between 25 and 30 study hours. However, the University of Helsinki grading scale does not fully correspond to the ECTS grading scale [13]. **Duration** is typically 12 weeks unless otherwise indicated. **Year** represents the last time the page was updated. In addition, Status implies that the course is still active as a module in a university program, otherwise, it has been discontinued with the material remaining online for the most part. Finally, **Other Courses** are university run courses which are either carried out as a once-off event, a summer school (The 2015 Natural Language Generation Summer School), a special tutorial or an intensive training course. Their status is marked as *archived*, since the material is still available online. As we can see from above, it appears that there are **only two** dedicated NLG university level courses on offer which appear to be active, while the remaining –upon close inspection– have been discontinued. Table 2 aligns the previously mentioned courses with assessment types, student level and provides some notes on course topics, and software tools mentioned. It also describes how research is linked to Teaching and Learning, so for instance *research-led*, implies that the course is instructional and teacher-led, and the module leader emphasises research content into the curriculum i.e., **NLG-1**. On the other hand, *Research-tutored*, also has a strong focus on research content, but it is student-led, with students selecting and presenting a research paper, followed by a discussion i.e., **NLG-3** [21]. With respect to assessment, the majority is a combination of continuous assessment (lab assignments) and a final (summative) exam or a single summative assignment, with the exception of **NLG-5**, which is project-based. As a result of our initial data collection, we can make the following observations:

- Our initial search has found **only two** dedicated university modules on Natural Language Generation that are currently active, which is of concern.
- The number of NLG dedicated modules available at university level is extremely small and has rather declined in recent years.
- The majority of courses teach or have taught Knowledge-based NLG (Rules, Grammars, or linguistic informed XML templates) with possibly some statistical NLG but less ML.
- While **NLG-2** and **NLG-5** cover newer ML methods for NLG, they do not appear to offer practical experience in ML techniques for building NLG components.
- Both the surface realisers – SimpleNLG [20] and OpenCCG [12] are the only NLG tools offered for teaching and learning.
- With the exception of the Scuba[14] corpus in **NLG01** and **NLG05**, we could not find any evidence suggesting the usage of any other NLG datasets.

[12]See http://ec.europa.eu/dgs/education_culture/repository/education/tools/docs/ects-guide_en.pdf

[13]https://guide.student.helsinki.fi/en/article/grades-and-assessment

4.2 Conceptualising NLG Research Themes

In this section, we present the outcome of a manual content analysis of the most recent systematic review for NLG [19]. The goal here is to extract a high level conceptual structure of the most up-to-date common research themes in NLG. The conceptual structure, which is effectively a classification tree (referenced as **NLG-CS-Research**) in Figure 1, reflects very closely the section and subsection headings in [19]. The reason for this is that it became evident when revising the tree that classifying deeper would result in capturing individual units of research (i.e., instances of publications and tools), which is out of scope. However, feedback from course content analysis did require modifying the tree to expand the NLG architectures concept deeper. We referred to [18] to include other types of architectures. In addition, when analysing the module **NLG-1**, it became evident that the concept **Modular Approaches** to NLG, in this context, aligned to rule-based tools, in particular, "plug and play" approaches that implement text planning as text schemas directly in Java [20] or well formed XML templates (**NLG-3**, **NLG-4**) [34], as opposed to a grammar-based formalism, such as [25] or [26]. We refer the reader to [19] for a detailed understanding of each concept in the classification tree. The tree is not meant to be exhaustive or authoritative, but in this research context, it is fit for purpose.

4.3 Aligning NLG University Course Content with Research

In Table 3, we align the subject matter content of the modules listed in Table 1, with the research concepts captured in the **NLG-CS-Research** classification tree in Figure 1. Based on this alignment, we can make the following observations:

- As one would expect, all NLG tasks are covered by each course.
- Teaching and Learning in NLG is dominated by the traditional pipeline/modular NLG architecture as described in [28], but there is little room for alternative architectures.
- Modular/Rule-based approaches to NLG tend to dominate within instructional (teacher-led) material.
- The evaluation of NLG systems has been covered in 50% of the courses, noting that **NLG-1** is the only active course with addresses the topic of evaluation.
- While stochastic approaches to planning have been catered for in the past, **NLG-5** (The 2005 NLG summer school) appears to be the only course which covers data-driven topics using an an instructional (teacher-led) approach, with the exception of deep learning approaches. However, we note the project-based assessment in **NLG-5** is rule-based with some language modelling.

[14]https://github.com/rdeoliveira/scuba/tree/master/src/nlg

Table 1: University Modules focuses only on NLG

University Courses							
Course	**Institution**	**Discipline**	**URL**	**ECTS**	**Duration**	**Year**	**Status**
NLG-1	University of Aberdeen	Computer Science	[9]	7.5	12 Weeks	2018	Active
NLG-3	University of Helsinki	Computer Science	[6]	2	7 Weeks	2017	Active
NLG-2	University of Edinburgh	Computer Science	[8]	5	12 Weeks	2013	Inactive
NLG-4	Ohio State University	Linguistics	[4]	NA	1 Semester	2006	Inactive
Other Courses							
NLG-5	NLG Summer School	Computer Science	[7]	Unknown	5 days	2015	Archived
NLG-5	University of Helsinki/Tartu	Computer Science	[2]	Unknown	4 days	2003	Archived

Table 2: University Modules focuses only on NLG, summarised by Topic and Assessment Type

University Courses and Course Topic				
Course	**Institution**	**Level**	**Assessment**	**Comments**
NLG-1	University of Aberdeen	Postgraduate	2 Hours Lecturer 2 Hour Lab 75%Exam (Summative) 25% Continuous Assessment	Research-led Knowledge-based NLG Evaluating NLG **Tools Used**: SimpleNLG, Scuba NLG
NLG-2	University of Edinburgh	Undergraduate Year 4 Postgraduate	70% Exam (Summative) 30% Continuous Assessment	Research-led Evaluating NLG Knowledge-based NLG NLG as Parsing Statistical NLG **Tools Used:** OpenCCG
NLG-3	University of Helsinki	Postgraduate	2 Hour Lecturer Seminar Group 1 Summative Assessment (100%)	Applied to Journalism Research-tutored Knowledge-based NLG Data-driven NLG Deep Learning NLG **Tools Used:** NA
NLG-4	Ohio State University	Undergraduate Advanced Final Year Postgraduate	100% Continuous Assessment	Research-led Knowledge-based NLG Statistical NLG **Tools Used:** XML/XSLT OpenCCG
Other Courses				
NLG-5	NLG Summer School (Consortium)	Industry Academia Postgraduate Undergraduate	100% Project-based Assessment	Knowledge-based NLG Statistical NLG Machine Learning-based NLG Readability Dialogue Systems Evaluating NLG **Tools Used:** SimpleNLG, ScubaNLG
NLG-6	University of Helsinki University of Tartu	Postgraduate	100% Continuous Assessment	Research-led Knowledge-based NLG Template/XML-based NLG Speech Synthesis **Tools Used:** XML/XSLT OpenCCG JavaSpeech API JSML

- While **NLG03** covers all data-driven approaches including deep learning, it is important to note that it is not instructional (teacher-led), but is rather student-led (research-tutored [21]), followed by a group discussion on the paper. Practical lab work does not appear to be part of the learning outcomes of this module.

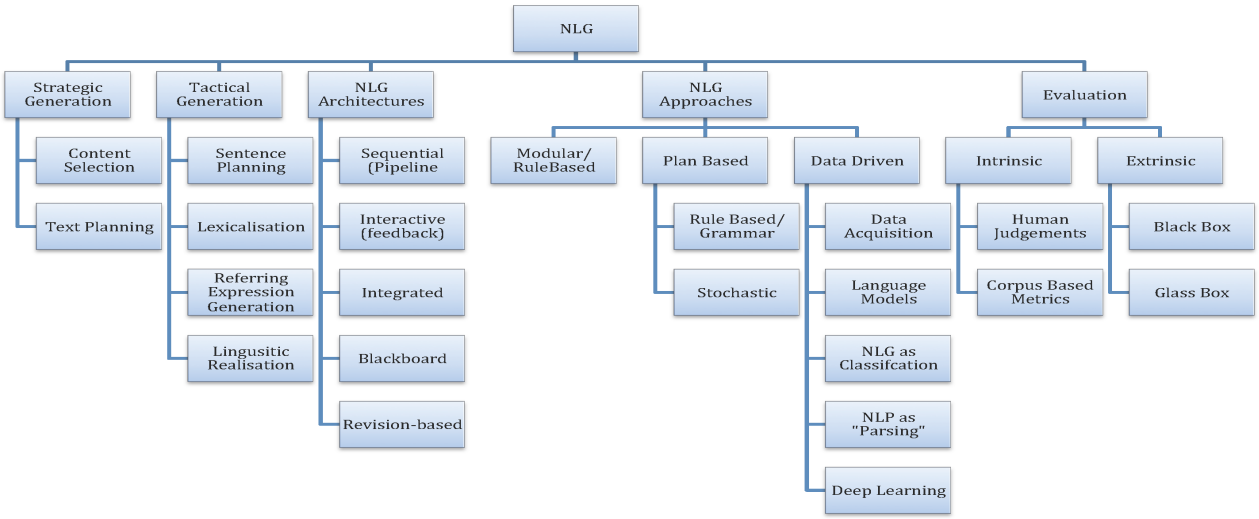

Figure 1: Classification Tree of Research Topics in NLG

			NLG Courses					
		Research Concepts	NLG-1	NLG-2	NLG-3	NLG-4	NLG-5	NLG-6
NLG Tasks	Strategic Generation	Content Section	Yes	Yes	Yes	Yes	Yes	Yes
		Text Planning	Yes	Yes	Yes	Yes	Yes	Yes
	Tactical Generation	Sentence Aggregation	Yes	Yes	Yes	Yes	Yes	Yes
		Lexicalisation	Yes	Yes	Yes	Yes	Yes	Yes
		Referring Expression Generation	Yes	Yes	Yes	Yes	Yes	Yes
		Linguistic/Surface Realisation	Yes	Yes	Yes	Yes	Yes	Yes
NLG Architectures		Sequential (Pipeline)	Yes	Yes	Yes	Yes	Yes	Yes
		Interactive (feedback)	No	No	Partially	No	No	No
		Integrated	No	No	No	No	No	No
		Blackboard	No	No	No	No	No	No
		Revision-based	No	No	No	No	No	No
NLG Approaches	Modular (Rule-based)		Yes	Yes	No	Yes	Yes	Yes
	Plan-based	Rule/Grammar-based	No	Yes	No	Yes	Yes	Yes
		Stochastic	No	Yes	No	Yes	No	Yes
	Data-driven	Data Acquisition	No	No	No	No	Yes	No
		Language Models	Yes	Yes	Yes	Yes	Yes	No
		NLG as Classification	No	No	Yes	No	Yes	No
		NLG as Parsing	No	No	Yes	No	Yes	No
		Deep Learning	No	No	Yes	No	No	No
Evaluation	Intrinsic	Human Judgement	Yes	Yes	No	No	Yes	No
		Corpus-based Metrics	Yes	Yes	No	No	Yes	No
	Extrinsic	Blackbox	Yes	Yes	No	No	Yes	No
		Glassbox	Yes	Yes	No	No	Yes	No

Table 3: NLG Research themes as concepts aligned to University Course Content

4.4 Limitations of the Study

The goal of this study was not to elicit knowledge at this time in order to affect any change in field. So the aim is not to elicit knowledge for critical evaluation by taking a negative standpoint. A course content review alone would not be sufficiently empirical to draw any concrete conclusions with respect to teaching and learning within the field [33]. In addition, attempting to evaluate teaching and learning activities (lesson plans and learning activities) in NLG is too broad a task. However, this study does provide knowledge as action which may inform efforts to bring about improvement to the prevailing practice within the NLG community [33]. The type of inquiry is conservative involving basic gap spotting and neglect spotting between NLG research content knowledge and content knowledge for NLG Teaching and Learning [29]. The study does not attempt to reject existing teaching practice or course design, but rather highlights possible room for improvement. Any critical stance would be inaccurate and not representative given the absence of a more rounded analysis containing additional qualitative and quantitative data, such as first had accounts, lecturer interviews and student evaluations, as opposed to just samples of course content.

In addition, this study does not include NLP university courses that provide one teaching session on NLG. Examples of this include the popular and new module on Natural Language Processing with Deep Learning by Stanford University[15], which covers the topic of abstractive summarisation. There are older NLP courses containing one teaching session dedicated to NLG which are still available, such as MIT OpenCourseware [16], modules from the University of Illinois[17] and an older inactive course on Text Production from the University of Bremen[18]. In addition, we excluded resources, such as a valuable guide on Deep Learning for NLG by the University of Stanford[19], as well as a specialist tutorial from the University of Cambridge[20] and a recent conference tutorial [21] presented at NAACL[22], June 1st, 2018, (USA). The justification is that none of these events are formal university courses.

4.5 Discussion

The data presented above confirms the high scarcity of university level courses in NLG, which is very concerning. Moreover, while industry demand for NLG has increased, the number of courses has over the years declined. Rule-based approaches and rule-based NLG tools including, templates and grammar-based formalisms with some language modelling tend to dominate. While newer material and some recent courses are beginning to provide knowledge transfer of machine learning approaches, including newer deep learning techniques, there does not appear to be much practical hands-on material available for developers to exploit beyond user generated content on the Web. A number of factors may be at play, including the lack of NLG tools and datasets suitable for teaching and learning, as well as institutional constraints with respect to demands on resources, which may have contributed to this deficit. Furthermore, in the past, a standalone NLG may not have been a priority within a computer science or computational linguistics curriculum due to a lack industry demand for NLG skills. However, this is speculative as a more rounded qualitative analysis would be needed to elicit this data.

One possible cost-effective solution, given the reasonable closeness of the community and limited resources, could involve a collaborative effort to generate course content material for a generic introductory course in NLG. The material, including learning activities, could be made available online, subject to agreed Intellectual Property (IP) and licensing issues. A *flipped* classroom could be applied [23], whereby students study the shared online content outside of classroom times and engage in practical learning activities in class, which could be local instances of the module led by one of the contributors at their home institution. Each individual module leader could add their own research-led knowledge and resources into the curriculum subject to shared agreement, so that the course is fully comprehensive and up-to-date. A further option would be to make a more streamlined version of the material available online as instance of a MOOC (Massive Open Online Course) [22].

5 CONCLUSION

As investment and demand for Artificial Intelligence technologies from industry continues to rise, so will the demand for training material, tools and know-how for building NLG systems. This study has analysed academic course content for the purpose of classifying and aligning NLG teaching and learning material with corresponding research concepts in the field. In conclusion, the university course offerings in general remain thin, and the availability of other non-university courses is even scarcer, which presents a challenge to a newcomer from industry seeking to demystify the usage of NLG tools and techniques, in particular newer deep learning approaches. Unless there is a significant investment in increased delivery of NLG courses at the university level involving an increase of knowledge and transfer of NLG skills and competencies, it is questionable whether the forecast industry demands for NLG technologies will be met.

[15] http://web.stanford.edu/class/cs224n/syllabus.html
[16] https://tinyurl.com/y954p9o7
[17] https://tinyurl.com/y9nxh4ge
[18] http://www.fb10.uni-bremen.de/anglistik/ling/ss02/nlg.htm
[19] https://cs.stanford.edu/~zxie/textgen.pdf
[20] https://tinyurl.com/yco9aatl
[21] https://tinyurl.com/ybd6r7pq
[22] http://naacl2018.org/

REFERENCES

[1] 2002. *ETMTNLP '02: Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics - Volume 1.* Association for Computational Linguistics, Stroudsburg, PA, USA.

[2] 2003. *XML-based Natural Language Generation.* http://www.ling.helsinki.fi/~gwilcock/Tartu-2003/

[3] 2005. *TeachNLP '05: Proceedings of the Second ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics.* Association for Computational Linguistics, Stroudsburg, PA, USA.

[4] 2006). *Introduction to Natural Language Generation Ling 795.10, Spring '06.* https://www.asc.ohio-state.edu/white.1240/teaching/SP-06/795.10/

[5] 2008. *TeachCL '08: Proceedings of the Third Workshop on Issues in Teaching Computational Linguistics.* Association for Computational Linguistics, Stroudsburg, PA, USA.

[6] 2013. *Spring 2011: Natural Language Generation.* http://www.inf.ed.ac.uk/teaching/courses/nlg/

[7] 2015). *Summer School on Natural Language Generation, Summarisation, and Dialogue Systems.* https://nlgsummer.github.io/

[8] 2017. *NATURAL LANGUAGE GENERATION FOR NEWS AUTOMATION.* https://courses.helsinki.fi/fi/582767/117339112

[9] 2018). *CS551H: NATURAL LANGUAGE GENERATION (2017-2018).* https://www.abdn.ac.uk/registry/courses/postgraduate/2018/computing_science/cs551h

[10] Stephen A Bartos, Norman G Lederman, and Judith S Lederman. 2014. Teachers' reflections on their subject matter knowledge structures and their influence on classroom practice. *School Science and Mathematics* 114, 3 (2014), 125–138.

[11] Steven Bird. 2008. Defining a Core Body of Knowledge for the Introductory Computational Linguistics Curriculum. In *Proceedings of the Third Workshop on Issues in Teaching Computational Linguistics (TeachCL '08)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 27–35. http://dl.acm.org/citation.cfm?id=1627306.1627311

[12] Cem Bozsahin, Geert-Jan M Kruijff, and Michael White. 2005. Specifying grammars for OpenCCG: A rough guide. *Included in the OpenCCG distribution* (2005).

[13] Lillian Boots Cassel, Andrew McGettrick, and Robert H Sloan. 2006. A comprehensive representation of the computing and information disciplines. In *ACM SIGCSE Bulletin*, Vol. 38. ACM, 199–200.

[14] Louis Cohen, Lawrence Manion, and Keith Morrison. 2002. *Research methods in education.* Routledge.

[15] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. 2011. Natural language processing (almost) from scratch. *Journal of Machine Learning Research* 12, Aug (2011), 2493–2537.

[16] Glynis Cousin. 2009. *Researching learning in higher education: An introduction to contemporary methods and approaches.* Routledge.

[17] Computing Curricula. 2001. Computer Science, Final Report, The Joint Task Force on Computing Curricula. *IEEE Computer Society and Association for Computing Machinery, IEEE Computer Society* (2001).

[18] Koenraad De Smedt, Helmut Horacek, and Michael Zock. 1996. Architectures for natural language generation: Problems and perspectives. In *Trends in Natural Language Generation An Artificial Intelligence Perspective*. Springer, 17–46.

[19] Albert Gatt and Emiel Krahmer. 2018. Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research* 61 (2018), 65–170.

[20] Albert Gatt and Ehud Reiter. 2009. SimpleNLG: A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation*. Association for Computational Linguistics, 90–93.

[21] Mick Healey. 2005. Linking research and teaching exploring disciplinary spaces and the role of inquiry-based learning. *Reshaping the university: New relationships between research, scholarship and teaching* (2005), 67–78.

[22] Andreas M Kaplan and Michael Haenlein. 2016. Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster. *Business Horizons* 59, 4 (2016), 441–450.

[23] Alison King. 1993. From sage on the stage to guide on the side. *College teaching* 41, 1 (1993), 30–35.

[24] Bertrand Meyer. 2006. Testable, reusable units of cognition. *Computer* 39, 4 (2006), 20–24.

[25] Johanna D Moore and Cécile L Paris. 1993. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational linguistics* 19, 4 (1993), 651–694.

[26] Crystal Nakatsu and Michael White. 2010. Generating with discourse combinatory categorial grammar. *Linguistic Issues in Language Technology* 4, 1 (2010), 1–62.

[27] Michela Pedroni, Manuel Oriol, and Bertrand Meyer. 2007. A framework for describing and comparing courses and curricula. In *ACM SIGCSE Bulletin*, Vol. 39. ACM, 131–135.

[28] Ehud Reiter and Robert Dale. 2000. *Building natural language generation systems.* Cambridge university press.

[29] Jörgen Sandberg and Mats Alvesson. 2011. Ways of constructing research questions: gap-spotting or problematization? *Organization* 18, 1 (2011), 23–44.

[30] Lee Shulman. 1987. Knowledge and teaching: Foundations of the new reform. *Harvard educational review* 57, 1 (1987), 1–23.

[31] Lee S Shulman. 1986. Those who understand: Knowledge growth in teaching. *Educational researcher* 15, 2 (1986), 4–14.

[32] Malcolm Tight. 2013. Discipline and methodology in higher education research. *Higher Education Research & Development* 32, 1 (2013), 136–151.

[33] Mike Wallace and Alison Wray. 2016. *Critical reading and writing for postgraduates.* Sage.

[34] Graham Wilcock. 2001. Pipelines, templates and transformations: XML for natural language generation. In *Proceedings of the 1st NLP and XML Workshop*. 1–8.