ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/336611715

Selected Challenges in Grammar-Based Text Generation from the Semantic Web

Chapter · October 2019

DOI: 10.1007/978-3-030-33274-7_5

CITATIONS
0

READS
45

1 author:

Simon Mille
University Pompeu Fabra

55 PUBLICATIONS **509** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project KRISTINA View project

Project PESCaDO View project

All content following this page was uploaded by Simon Mille on 07 January 2020.

The user has requested enhancement of the downloaded file.

Selected Challenges in Grammar-Based Text Generation from the Semantic Web

Simon Mille(✉)

Universitat Pompeu Fabra, Barcelona, Spain
simon.mille@upf.edu
http://www.springer.com/gp/computer-science/lncs

Abstract. In this paper, based on the recent outcome of two shared tasks on structured data verbalisation, and examining one system in particular, we present some evidence why grammar-based systems are particularly relevant for the verbalisation of structured data as found in the Semantic Web. We then define possible future lines of research, centered around the FORGe system and the linguistic grounding of Semantic Web datasets.

Keywords: Natural Language Generation · Semantic Web · Grammar-based systems

1 Introduction

Nowadays, thanks to Semantic Web (SW) initiatives such as the W3C Linking Open Data project[1], a tremendous amount of structured knowledge is publicly available in the form of language-independent triples. The Linked Open Data (LOD) cloud[2] currently contains over a thousand interlinked datasets (as, for instance, DBpedia[3] or Wikidata[4]), which cover a large range of domains and amount to billions of different triples. Table 1 shows three DBpedia triples –Subject, Property, Object- related to black rice (arròs negre), representing its origin and two of its ingredients.[5]

The LOD datasets are frequently enriched, manually or automatically, by extracting knowledge from, e.g., multimedia and textual data, and populate already available domain models (i.e., known class and property definitions), as well as acquire new –*unknown*– ones (ontology learning). Their formal knowledge representation allows for applying powerful algorithms which have proved crucial

[1] https://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData.
[2] http://lod-cloud.net/.
[3] https://wiki.dbpedia.org/.
[4] https://www.wikidata.org/wiki/Wikidata:Main_Page.
[5] This information appears in the infobox of the corresponding Wikipedia page: https://en.wikipedia.org/wiki/Arr%C3%B2s_negre.

© Springer Nature Switzerland AG 2019
G. S. Osipov et al. (Eds.): Artificial Intelligence, LNAI 11866, pp. 85–95, 2019.
https://doi.org/10.1007/978-3-030-33274-7_5

Table 1. Sample triples from the food domain (from the English DBpedia).

	Subject	Property	Object
Triple 1	Arròs_negre	Country	Spain
Triple 2	Arròs_negre	Ingredient	White_rice
Triple 3	Arròs_negre	Ingredient	Squid

in fields such as Question Answering [19]. However, there has been relatively little research in applying Natural Language Generation (NLG, or Text Generation) techniques to the multilingual verbalisation from Semantic Web contents, as exemplified in Table 2.

Table 2. Possible verbalisations of the triples from Table 1.

Language	Possible verbalisation
English 1	Arròs negre is a dish from Spain. It contains white rice and squid
English 2	White rice and squid are basic ingredients of the Spanish dish called Arròs negre
Spanish 1	Arròs negre es un plato español basado en arroz blanco y calamar

Existing NLG systems for Semantic Web contents are, on the one hand, not adapted to the richness nor the constant evolution of the LOD cloud and its target users, and, on the other hand, application- and language-specific, and thus have coverage and/or reliability issues. As a result, their usefulness, robustness and portability –some of the most challenging issues in NLG nowadays [8]– are limited. In the remainder of the paper, we argue that grammar-based NLG is efficient for text generation from LOD data; we give a short overview of the most widely used NLG techniques for generating texts from structured data, point out their current limitations, and define a few aspects that would need to be developed in the near future in order to ensure scalability and reusability.

2 Natural Language Generation in the Context of the Semantic Web

2.1 Approaches and Limitations

The Semantic Web and Natural Language Generation communities have for a long time been disconnected: one of the primary applications of the Semantic Web resources is Question Answering, for which the understanding of the questions and the retrieval of the answers is the main focus, rather than the verbalisation of the triples; indeed, returning a simple list of triples as answer to a question may suffice [17]. Very recently, in 2015, the first International Workshop on Natural Language Generation from the Semantic Web was organised in France (WebNLG[6]). As mentioned in the workshop call of papers, on the one hand, Semantic Web applications need to make the contents accessible the potential users, and on the other hand, NLG-based approaches have been used for verbalising structured data coming from, e.g., ontologies [5,13] or time series [3,7]. These two areas are complementary but relatively few attempts have been made to bring them together.

Traditionally, Natural Language Generation is viewed a sequence of three subtasks: (i) content selection, which is responsible for determining the contents to be rendered as text, (ii) text planning, which takes care of packaging the contents into discursively organised units (i.e., sentences) and (iii) linguistic generation, which realises the contents as well-formed text [30]. The advantage of splitting language generation into specific tasks is to allow for a precise and independent modelling of each level of language description (semantics, syntax, topology, morphology).

For the verbalisation of structured data, there are three main approaches to realise each of these subtasks [8,16]: (i) filling slot values in predefined sentence templates (e.g. [1]), (ii) applying grammars (rules) that encode different types of linguistic knowledge (e.g. [35]), and (iii) predicting statistically the most appropriate output (e.g. [4,15]). Template-based systems are very reliable in terms quality, but are the worst in terms of portability since new templates need to be defined for every new domain, style, language, etc. Statistical systems have the widest coverage, but the relevance and the quality of the produced texts cannot be ensured. Furthermore, they are fully dependent on the available –scarce and mostly monolingual– training data [23]. The development of grammar-based systems is time-consuming and usually they have coverage issues, but they are easy to port to a new domain (and also style, language, etc.), do not require training material, allow for a greater control over the outputs (e.g. for mitigating possible errors or tuning the output to a desired style), and the linguistic knowledge used for one domain or language can be reused for other domains and languages. However, due to their complexity, such approaches have undergone few developments within the open-source community in the recent years [16]. In addition, a number of systems actually address the whole sequence as one step, by combining approaches (i) and (iii) and filling the slot values of pre-existing templates using neural network techniques [26].

Last but not least, let us note that another limitation to the Natural Language Generation from Semantic Web data is that until recently, NLG from SW data has been applied to independent datasets only, leaving aside multiple interlinked datasets,[7] and a large part of it has focused on the description of the knowledge model rather than on the verbalisation of the contents [33].

---

[6] http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=45093.

[7] One of the first papers mentioning multiple datasets was published in 2012 [10].

2.2 FORGe: An Example of a Grammar-Based System

FORGe is an open-source generator developed at the Pompeu Fabra University, implemented as graph-transducers and that covers the last two NLG subtasks (text planning and linguistic generation). FORGe, according to the lines of the Meaning-Text Theory [21], is based on the notion of linguistic dependencies, that is, the semantic, syntactic and morphologic relations between the components of the sentence. Input predicate-argument structures are mapped onto sentences by applying a series of rule-based graph-transducers. The generator handles Semantic Web (SW) inputs by means of introducing abstract predicate-argument (PredArg) templates and micro-planning grammars as previous steps to the core linguistic generation module. A sample PredArg template is shown in Fig. 1: the DBpedia property *floorArea* is mapped to the predicate *floor_area*, which has two arguments, a building and a surface area, which are respectively the subject and the object of the property. Lists of such instantiated (*populated*) PredArg structures are passed on to the generator.

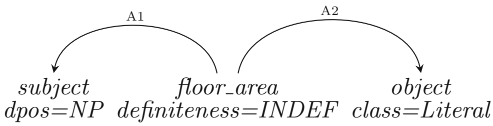

Fig. 1. Sample PredArg template for the floor area of a building

For micro planning, on the one hand, generic rules look for shared pairs of predicate and subject argument in the populated templates, and introduce coordinations or quasi-coordinations between the two objects as in: *[Alan Bean]*$_S$ *[was born]*$_P$ *[in Wheeler]*$_{O1}$ *[on March 15, 1932]*$_{O2}$. Other generic rules check if an argument of a predicate appears further down in the ordered list of PredArg structures. If so, the PredArg structures are merged by fusing the common argument; during linguistic generation, this results in the introduction of post-nominal modifiers such as relative and participial clauses or appositions; e.g. *250 Delaware Avenue*$_S$, *which has a [floor area]*$_{P2}$ *of [30843.8 square meters]*$_{O2}$, *[is]*$_{P1}$ *in Buffalo*$_{O1}$). On the other hand, rules specific to a domain (e.g restaurant domain here) have been implemented so as to aggregate objects that have not been aggregated by the generic rules; between the brackets, detail of the restrictions about co-occurring properties:

- eatType$_{P1}$ + priceRange$_{P2}$: *a cheap*$_{O2}$ *pub*$_{O1}$.
- eatType$_{P1}$ + familyFriendly$_{P2}$: *a family-friendly*$_{O2}$ *restaurant*$_{O1}$.
- etc.

For rendering of the aggregated PredArg structures into sentences, the core FORGe grammars [24] perform the following actions: (i) syntacticisation of predicate-argument graphs; (ii) introduction of function words; (iii) linearisation and retrieval of surface forms. First, a deep-syntactic structure is generated: missing parts of speech are assigned, the syntactic root of the sentence is chosen, and from there, a syntactic tree over content words is built node by node. Then, idiosyncratic words (prepositions, auxiliaries, determiners, etc.) are introduced and fine-grained (surface-)syntactic labels are established, using a subcategorisation lexicon. For this purpose, lexical resources are used that can be derived from PropBank [18] or VerbNet [31]; see [25]. Personal and relative pronouns are introduced using the *class* feature, which allows for distinguishing between human and non-human antecedents. Finally, morpho-syntactic agreements are resolved, the syntactic tree is linearised, through the ordering of (i) governor/dependent and (ii) dependents with each other, and the surface forms are retrieved. Post-processing rules are then applied: upper casing, replacement of underscores by spaces, etc.

For illustrating these three steps, consider the *floorArea* property of Fig. 1, from the WebNLG dataset, and selected phenomena: (i) the support verb *be* is established as the root, (ii) the preposition *of* is introduced, and (iii) the *subject* relation between *be* and *floor area* causes the former to be placed after the latter and get morphological agreement features from it (third person singular): *the floor* $area_{3sg}$ *of building* $X > is_{3sg}\ N\ m^2$.

3 What Grammar-Based Systems Are Good For: Lessons Learnt from the WebNLG and E2E Challenges

3.1 The WebNLG and E2E Challenges

In the past two years, two NLG challenges starting from structured data took place, namely WebNLG [14] and E2E [27], and different types of systems were used to produce outputs: template-based, rule-based, statistical machine translation-based, recurrent neural network-based, etc.; see respective overview papers.

In the framework of the WebNLG challenge, the task consisted in generating texts from up to 7 DBpedia triples from 15 different categories, covering in total 373 different DBpedia properties. 9 categories appeared in the training data (Astronaut, Building, University, Monument, ComicsCharacter, Food, Airport, SportsTeam and WrittenWork), and six categories were “unseen”, in that they did not appear in the training data (Athlete, Artist, City, MeanOfTransportation, CelestialBody, and Politician). At the time of the challenge, the WebNLG dataset contained about 25K data-text pairs for about 10K distinct inputs, that is, about 2.5 reference sentences per triple set. The challenge thus focused on verbalising a wide range of inputs.

The input for the E2E challenge is very similar to the WebNLG challenge in the sense that it consisted of a list of up to 8 triples, corresponding to 8 properties from the Restaurant domain (name, location, nearby restaurants, type of food, type of restaurant, price range, customer rating, kid friendliness). The E2E data consisted of about 50K data-text pairs; only 108 different combinations of properties are found in the training set, which gives an average of about 500 reference sentences per triple set. Thus, unlike WebNLG, E2E focuses on the (many) different ways to render a specific set of properties. The evaluation data does not contain unseen combinations of properties. For both challenges, the datasets were released in triple format (*Subject*, *Property*, *Object*). In the following, another DBpedia triple and a target reference sentence are shown:

320_South_Boston_Building ∥ *architect* ∥ *George_Winkler*
George Winkler designed the 320 South Boston Building.

For both challenges, automatic and human evaluations were carried out; only the common metrics are reported here. BLEU [28] and METEOR [2] are *n*-gram-based metrics that compare word-to-word a candidate with a reference sentence. BLEU matches exact words only, whereas METEOR matches also synonyms; 100 and 1 respectively indicate sameness. For human evaluations, judges are asked to either rank or rate candidate sentences in terms of their adequacy with the input (*Semantics*), their linguistic correctness (*Grammar*), and their *Fluency*. Nine systems were evaluated for WebNLG, and twenty-one for E2E, and most competing systems follow statistical approaches.

3.2 Results and Discussion

Table 3 shows the results obtained by the FORGe system (see Sect. 2.2). The results of the two shared tasks were informative in different aspects. First of all, the best systems (neural systems) score much better than a rule-based system such as FORGe according to basic automatic evaluation metrics: for the BLEU metric, the difference is of about 20 points on seen data (WebNLG: 60.59/40.88, E2E 68.05/42.07), and FORGe obtains core among the lowest. However, when synonymy is taken into account, as in the METEOR metric, the gap is much smaller, and the system ranks can even sometimes be inverted (e.g,. WebNLG *seen*, in which FORGe ranks 3–4 instead of 8 according to BLEU). Second, the results of the human evaluations are rather different, with FORGe ranking much higher than according to the automatic evaluations: at WebNLG, FORGe is consistently in the first half of the ranking, and for E2E, in which the systems end up being clustered in five groups of equivalent ratings, it reaches the second cluster. In other words, even though the outputs produced by a grammar-based system do not reflect faithfully the reference outputs, they tend to be well accepted by human judges, in any case better accepted than suggested by the automatic metrics.

Finally, the evaluation on the unseen properties at WebNLG shows that FORGe was the most adaptable system, attaining the first rank according to both automatic and human metrics. When statistical systems have no training data, they are simply not able to generate correct texts, but a grammar-based system does not rely on the training data, and it is thus possible to tune it to new inputs at a reduced cost.[8] Note however that when enough good quality data is

[8] It took about two hours to adapt FORGe to a hundred new DBpedia properties.

Table 3. Scores (and rankings) of FORGe according to BLEU and METEOR, and rankings according to human evaluations (Semantics, Grammar, Fluency).

Dataset	BLEU	METEOR	Semantics	Grammar	Fluency
WebNLG all	38.65 (3–4)	0.39 (1)	1	1	1–2
WebNLG seen	40.88 (8)	0.40 (3–4)	2–4	1–2	4–5
WebNLG uns.	35.70 (1)	0.37 (1)	1	1	1
E2E all (seen)	42.07 (20)	0.37 (20)	10–14 (cluster 2/5)		

provided, the human evaluations of some neural systems can be astounding and even outperform human-written texts, as it was the case for the ADAPT system [11] on the WebNLG seen properties [32].

4 Towards an Efficient Verbalisation of Structured Data

4.1 Challenges for the FORGe Generator

In order to target a multilingual Natural Generation System, a large part of the core resources need to be language-independent and their portability to new languages need to be ensured, as well as their flexibility, so that the tool can be used not only in question/answering systems, but also in applications with more complex input structures that proceed from analysed text, such as text summarisation, text simplification and dialogue interfaces.

With respect to the development of graph-transduction grammars, three critical issues currently need to be highlighted: (i) the packaging of the selected triples into coherent groups (micro-planning), (ii) the definition of a valid sentence structure over each group of triples, and (iii) the creation of grammar-compatible NLG-oriented lexical resources. For the first subtask, micro-planning grammars need to be improved in the sense that they should allow for more types of aggregation than the one described in Sect. 2.2, and be made as domain-independent as possible. This can be done by looking at large amounts of textual data in order to compile the common aggregation patterns in the different languages and domains, and adapting the rules accordingly. For the second subtask, the sentence structuring module needs to be tested on a large scale in order to obtain wide coverage and ensure flexibility and multilinguality. Most of the rules at this level are generic and produce complete syntactic structures, but there is no validation of their correctness so far. As the input structures get more complex (i.e. as the number of triples to verbalize increases), defining a correct sentence structures can get more challenging, and here again the compilation of language-specific preferred syntactic patterns would help controlling the process. For the third subtask, the automatic extraction of NLG- and dependency-suited information from lexica and annotated data (e.g. VerbNet, PropBank) needs to be investigated. In particular, the syntactic information about the participants (e.g. if a preposition is needed –*ingredient of*) is not expressed directly in the existing resources. However, this subcategorisation information can be derived from, e.g., VerbNet, which is neither NLG– nor dependency-friendly. Some steps in this direction have already been done in [20,25].

4.2 Challenges in the Linguistic Grounding of LOD Datasets

As far as input representations are concerned, an NLG pipeline needs to be fed with linguistic structures. These are quite different from the triples found on the LOD cloud, in which the Properties are labeled with an open vocabulary and only two types of relations (Subject and Object) are used. The triples should be mapped onto linguistic concepts and relations, preferably according to standard lexico-semantic resources to ensure reusability (e.g. VerbNet, NomBank [22] and PropBank, which, thanks to the amount of multilingual resources connected to them, can be used as interlingua). For instance, the property *ingredient* as seen in Table 1 would need to be mapped to, e.g., the PropBank predicate contain.01, or the NomBank predicate ingredient.01, and the Subject and Object to be mapped to the corresponding participant slot according to their respective subcategorisation frames. Participant slots can be simple predicate-argument relations –first argument, second argument, etc.– or more "conceptual" relations such as Agent, Patient, Beneficiary. For an informed generation process, basic properties of the participants, e.g. classes such as their type (Country, Ingredient, Person) and gender (Female or Male), found on the LOD cloud in the form of respective class and property assertions, need to be added to the linguistic representation.

To the best of our knowledge, little research has been carried out so far on bringing together SW contents and standard linguistic resources in the context of NLG: on the one hand, standard SW approaches such as *lemon* [34] or word embeddings [29] define their own lexicons to be associated with the properties, and on the other hand linguistic resources such as VerbNet, NomBank and PropBank are generally not connected with SW knowledge bases. One exception is the PreMOn model [9], which specifically aims at linking VerbNet, NomBank, PropBank and FrameNet [12] with an ontology that models semantic classes and their roles; however, PreMon leaves open the mapping to specific LOD datasets such as DBpedia.[9]

The linguistic grounding challenge is thus focused on the mapping of triples onto abstract linguistic structures to serve as input for the NLG pipeline; this includes the mapping of (both *known* and *unknown*) properties, as well as of their arguments (Subject and Object), onto minimal language-independent linguistic structures that contain all the information needed for being verbalised. There is a need to innovate according to the linguistic grounding of database sub-structures based on class and property statements onto sentential semantic structures. Linguistically motivated interface structures based on PropBank and/or VerbNet representations need to be defined, as well as the mappings from both known and unknown LOD triples, to account for the high degree of dynamism of SW databases. The connection of SW lexicons with language-oriented ones need to

[9] See also [6] for an overview of models to represent linked data and their issues.

be investigated, together with the use of the semantics of the underlying schema (e.g. taxonomical information) in order to derive basic features (e.g. class, gender, etc.) or generalising the concepts in case some contents are very specific and cannot be mapped onto the exact concept.

5 Conclusions

Statistical text generation systems have been the main focus of the Natural Language Generation community in the past years. However, their low portability to new languages and domains and the lack of control over the final output, as well as the very limited amount of actual linguistic knowledge used during the generation process are currently obstacles to the widespread use of such systems on Semantic Web structured data. In this paper, we show how grammar-based systems are suitable for the verbalisation of structured data and discuss open challenges and future lines of research, centered around (i) the increase of both grammatical and lexical coverage and (ii) the linguistic grounding of Semantic Web datasets.

Acknowledgements. The work reported in this paper has been partly supported by the European Commission in the framework of the H2020 Programme under the contract numbers 700475-IA, 700024-RIA, 779962-RIA, 786731-RIA and 825079-ICT-STARTS.

References

1. Androutsopoulos, I., Lampouras, G., Galanis, D.: Generating natural language descriptions from OWL ontologies: the naturalowl system. J. Artif. Intell. Res. **48**, 671–715 (2013)
2. Banerjee, S., Lavie, A.: METEOR: an automatic metric for MT evaluation with improved correlation with human judgments. In: Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization, pp. 65–72 (2005)
3. Belz, A.: Automatic generation of weather forecast texts using comprehensive probabilistic generation-space models. J. Nat. Lang. Eng. **14**(4), 431–455 (2008)
4. Belz, A., White, M., Espinosa, D., Kow, E., Hogan, D., Stent, A.: The first surface realisation shared task: overview and evaluation results. In: Proceedings of the Generation Challenges Session at the 13th European Workshop on Natural Language Generation (ENLG), Nancy, France, pp. 217–226 (2011)
5. Bontcheva, K., Wilks, Y.: Automatic report generation from ontologies: The MIAKT approach. In: Meziane, F., Métais, E. (eds.) NLDB 2004. LNCS, vol. 3136, pp. 324–335. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-27779-8_28
6. Bosque-Gil, J., Gracia, J., Montiel-Ponsoda, E., Gómez-Pérez, A.: Models to represent linguistic linked data. Nat. Lang. Eng. **24**(6), 811–859 (2018)
7. Bouayad-Agha, N., Casamayor, G., Mille, S., Wanner, L.: Perspective-oriented generation of football match summaries: old tasks, new challenges. ACM Trans. Speech Lang. Process. **9**(2), 3:1–3:31 (2012)

8. Bouayad-Agha, N., Casamayor, G., Wanner, L.: Natural language generation in the context of the semantic web. Semant. Web **5**(6), 493–513 (2014)
9. Corcoglioniti, F., Rospocher, M., Aprosio, A.P., Tonelli, S.: PreMON: a lemon extension for exposing predicate models as linked data. In: Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC), pp. 877–884 (2016)
10. Dannélls, D., Damova, M., Enache, R., Chechev, M.: Multilingual online generation from semantic web ontologies. In: Proceedings of the 21st International Conference on World Wide Web, pp. 239–242. ACM (2012)
11. Elder, H., Gehrmann, S., O'Connor, A., Liu, Q.: E2E NLG challenge submission: towards controllable generation of diverse natural language. In: Proceedings of the 11th International Conference on Natural Language Generation, pp. 457–462 (2018)
12. Fillmore, C.J., Baker, C.F., Sato, H.: The FrameNet database and software tools. In: Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC), Las Palmas, Canary Islands, Spain, pp. 1157–1160 (2002)
13. Galanis, D., Androutsopoulos, I.: Generating multilingual descriptions from linguistically annotated OWL ontologies: the naturalowl system. In: Proceedings of the Eleventh European Workshop on Natural Language Generation, pp. 143–146. Association for Computational Linguistics (2007)
14. Gardent, C., Shimorina, A., Narayan, S., Perez-Beltrachini, L.: Creating training corpora for micro-planners. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, Vancouver, Canada, August 2017
15. Gardent, C., Shimorina, A., Narayan, S., Perez-Beltrachini, L.: The WebNLG challenge: generating text from RDF data. In: Proceedings of the 10th International Conference on Natural Language Generation, pp. 124–133 (2017)
16. Gatt, A., Krahmer, E.: Survey of the state of the art in natural language generation: core tasks, applications and evaluation. J. Artif. Intell. Res. **61**, 65–170 (2018)
17. Höffner, K., Walter, S., Marx, E., Usbeck, R., Lehmann, J., Ngonga Ngomo, A.C.: Survey on challenges of question answering in the semantic web. Semant. Web **8**(6), 895–920 (2017)
18. Kingsbury, P., Palmer, M.: From TreeBank to PropBank. In: Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC), Las Palmas, Canary Islands, Spain, pp. 1989–1993 (2002)
19. Kwok, C., Etzioni, O., Weld, D.S.: Scaling question answering to the web. ACM Trans. Inf. Syst. (TOIS) **19**(3), 242–262 (2001)
20. Lareau, F., Lambrey, F., Dubinskaite, I., Galarreta-Piquette, D., Nejat, M.: GenDR: a generic deep realizer with complex lexicalization. In: Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC), Miyazaki, Japan, pp. 3018–3025 (2018)
21. Mel'čuk, I.: Dependency Syntax: Theory and Practice. State University of New York Press, Albany (1988)
22. Meyers, A., et al.: The NomBank project: an interim report. In: Proceedings of the Workshop on Frontiers in Corpus Annotation, Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT/NAACL), Boston, MA, USA, pp. 24–31 (2004)

23. Mille, S., Belz, A., Bohnet, B., Graham, Y., Pitler, E., Wanner, L.: The first multilingual surface realisation shared task (SR 2018): overview and evaluation results. In: Proceedings of the 1st Workshop on Multilingual Surface Realisation (MSR), 56th Annual Meeting of the Association for Computational Linguistics (ACL), Melbourne, Australia, pp. 1–12 (2018)
24. Mille, S., Carlini, R., Burga, A., Wanner, L.: FORGe at SemEval-2017 task 9: deep sentence generation based on a sequence of graph transducers. In: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), Vancouver, Canada, pp. 917–920. Association for Computational Linguistics, August 2017. http://www.aclweb.org/anthology/S17-2158
25. Mille, S., Wanner, L.: Towards large-coverage detailed lexical resources for data-to-text generation. In: Proceedings of the First International Workshop on Data-to-text Generation, Edinburgh, Scotland (2015)
26. Nayak, N., Hakkani-Tür, D., Walker, M.A., Heck, L.P.: To plan or not to plan? discourse planning in slot-value informed sequence to sequence models for language generation. In: Proceedings of INTERSPEECH, Stockholm, Sweden, pp. 3339–3343 (2017)
27. Novikova, J., Dušek, O., Rieser, V.: The E2E dataset: new challenges for end-to-end generation. In: Proceedings of the 18th Annual Meeting of the Special Interest Group on Discourse and Dialogue, Saarbrücken, Germany (2017). https://arxiv.org/abs/1706.09254, arXiv:1706.09254
28. Papineni, K., Roukos, S., Ward, T., Zhu, W.J.: BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, pp. 311–318. Association for Computational Linguistics (2002)
29. Perez-Beltrachini, L., Gardent, C.: Learning embeddings to lexicalise RDF properties. In: * SEM 2016, The Fifth Joint Conference on Lexical and Computational Semantics, pp. 219–228 (2016)
30. Rambow, O., Korelsky, T.: Applied text generation. In: Proceedings of the 3rd Conference on Applied Natural Language Processing (ANLP), Trento, Italy, pp. 40–47 (1992)
31. Schuler, K.K.: VerbNet: a broad-coverage, comprehensive verb lexicon. Ph.D. thesis, University of Pennsylvania (2005)
32. Shimorina, A., Gardent, C., Narayan, S., Perez-Beltrachini, L.: The WebNLG challenge: report on human evaluation. Technical report, Université de Lorraine, Nancy, France (2017)
33. Stevens, R., Malone, J., Williams, S., Power, R., Third, A.: Automating generation of textual class definitions from OWL to English. J. Biomed. Semant. **2**, S5 (2011). BioMed Central
34. Walter, S., Unger, C., Cimiano, P.: M-ATOLL: a framework for the lexicalization of ontologies in multiple languages. In: Mika, P., et al. (eds.) ISWC 2014. LNCS, vol. 8796, pp. 472–486. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11964-9_30
35. Wanner, L., Bohnet, B., Bouayad-Agha, N., Lareau, F., Nicklaß, D.: MARQUIS: generation of user-tailored multilingual air quality bulletins. Appl. Artif. Intell. **24**(10), 914–952 (2010)

View publication stats