



HAL
open science

CLEF Workshop JOKER: Automatic Wordplay and Humour Translation

Liana Ermakova, Tristan Miller, Orlane Puchalski, Fabio Regattin, Élise Mathurin, Sílvia Araújo, Anne-Gwenn Bosser, Claudine Borg, Monika Bokinić, Gaelle Le Corre, et al.

► **To cite this version:**

Liana Ermakova, Tristan Miller, Orlane Puchalski, Fabio Regattin, Élise Mathurin, et al.. CLEF Workshop JOKER: Automatic Wordplay and Humour Translation. *Advances in Information Retrieval*, 13186, Springer International Publishing, pp.355-363, 2022, Lecture Notes in Computer Science, 10.1007/978-3-030-99739-7_45 . hal-03637773

HAL Id: hal-03637773

<https://hal.science/hal-03637773>

Submitted on 11 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CLEF Workshop JOKER: Automatic Wordplay and Humour Translation

Liana Ermakova¹[0000-0002-7598-7474], Tristan Miller²[0000-0002-0749-1100],
Orlane Puchalski¹, Fabio Regattin³[0000-0003-3000-3360], Élise Mathurin¹, Silvia
Araújo⁴[0000-0003-4321-4511], Anne-Gwenn Bosser⁵, Claudine
Borg⁶[0000-0003-3858-5502], Monika Bokinić⁷, Gaelle Le Corre⁸, Benoît
Jeanjean¹, Radia Hannachi⁹, Gorg Mallia⁶, Gordan Matas¹⁰, and Mohamed
Saki¹

¹ Université de Bretagne Occidentale, HCTI EA-4249/MSHB, 29200 Brest, France

² Austrian Research Institute for Artificial Intelligence, Vienna, Austria

³ Dipartimento DILL, Università degli Studi di Udine, 33100 Udine, Italy

⁴ Universidade do Minho, CEHUM, Rua da Universidade, 4710-057 Braga, Portugal

⁵ École Nationale d'Ingénieurs de Brest, Lab-STICC CNRS UMR 6285, France

⁶ University of Malta, Msida MSD 2020, Malta

⁷ University of Gdansk, Poland

⁸ Université de Bretagne Occidentale, CRBC, 29200 Brest, France

⁹ Université de Bretagne Sud, HCTI EA-4249, 56321 Lorient, France

¹⁰ University of Split, Croatia

Abstract Humour remains one of the most difficult aspects of intercultural communication: understanding humour often requires understanding implicit cultural references and/or double meanings, and this raises the question of its (un)translatability. Wordplay is a common source of humour in due to its attention-getting and subversive character. The translation of humour and wordplay is therefore in high demand. Modern translation depends heavily on technological aids, yet few works have treated the automation of humour and wordplay translation, or the creation of humour corpora. The goal of the JOKER workshop is to bring together translators and computer scientists to work on an evaluation framework for wordplay, including data and metric development, and to foster work on automatic methods for wordplay translation. We propose three pilot tasks: (1) classify and explain instances of wordplay, (2) translate single words containing wordplay, and (3) translate entire phrases containing wordplay.

Keywords: Machine translation · humour · wordplay · puns · parallel corpora · evaluation metrics · creative language analysis

1 Introduction

Intercultural communication relies heavily on translation. Humour remains by far one of its most difficult aspects; to understand humour, one often has to grasp implicit cultural references and/or capture double meanings, which of

course raises the question of the (un)translatability of humour. One of the most common sources of humour is wordplay, which involves the creative application or bending of rules governing word formation, choice, or usage. Wordplay is used by novelists, poets, playwrights, scriptwriters, and copywriters, and is often employed in titles, headlines, proper nouns, and slogans for its ability to grab attention and for its mnemonic, playful, or subversive character. The translation of wordplay is therefore in high demand. But while modern translation is heavily aided by technological tools, virtually none has any specific support for humour and wordplay, and there has been very little research on the automation of humour and wordplay translation. Furthermore, most AI-based translation tools require a quality and quantity of training data (e.g., parallel corpora) that has historically been lacking of humour and wordplay.

Preserving wordplay can be crucial for maintaining the pragmatic force of discourse. Consider the following pun from *Alice’s Adventures in Wonderland* by Lewis Carroll, which exploits the homophony of *lessons* and *lessens* for humorous effect: “‘That’s the reason they’re called lessons,’ the Gryphon remarked: ‘because they lessen from day to day.’” Henri Parisot’s French translation manages to preserve both the sound and meaning correspondence by using the pair *cours/courts*: “*C’est pour cette raison qu’on les appelle des cours : parce qu’ils deviennent chaque jour un peu plus courts.*” By contrast, Google Translate uses the pair *leçons/diminuent* and the sentence becomes nonsensical: “*‘C’est la raison pour laquelle on les appelle leçons, remarqua le Griffon : parce qu’elles diminuent de jour en jour.’*”

The goal of the JOKER workshop is to bring together translators, linguists and computer scientists to work on an evaluation framework for creative language. All types of contributions will be welcomed: this includes research, survey, position, discussion, and demo papers, as well as extended abstracts of published papers. We will also oversee pilot tasks making use of a new, multilingual parallel corpus of wordplay and humour that we have produced: **Pilot Task 1** is to classify single words containing wordplay according to a given typology, and provide lexical-semantic interpretations; **Pilot Task 2** is to translate single words containing wordplay; and **Pilot Task 3** is to translate entire phrases that subsume or contain wordplay. The two translation tasks will initially target English and French but may be expanded to further languages as data becomes available. As discussed in §3 below, the consideration of appropriate evaluation metrics for these tasks is one of the goals of the workshop.

To encourage participants to use our data in creative ways and to collect ideas for future editions of the workshop, we also propose an **Unshared Task**. We particularly welcome ideas from researchers in the humanities on how we can promote deeper linguistic and social-scientific analysis of our data.

2 Background

Automatic Humour Analysis and Related Campaigns. To date, there have been only a handful of studies on the machine translation (MT) of wordplay. Farwell

and Helmreich [7] proposed a pragmatics-based approach to MT that accounts for the author’s locutionary, illocutionary, and perlocutionary intents (that is, the “how”, “what”, and “why” of the text), and discuss how it might be applied to puns. However, no working system appears to have been implemented. Miller [19] proposed an interactive method for the computer-assisted translation of puns, an implementation and evaluation of which was described by Kolb and Miller [16]. Their study was limited to a single language pair (English to German) and translation strategy (namely, the PUN→PUN strategy described below).

Numerous studies have been conducted for the related tasks of humour generation and detection. Pun generation systems have often been based on template approaches. Valitutti and al. [26] used lexical constraints to generate adult humour by substituting one word in a pre-existing text. Hong and Ong [11] trained a system to automatically extract humorous templates which were then used for pun generation. Some current efforts to tackle this difficult problem more generally using neural approaches have been hindered by the lack of a sizable pun corpus [29]. Meanwhile, the recent rise of conversational agents and the need to process large volumes of social media content point to the necessity of automatic humour recognition [21]. Humour and irony studies are now crucial when it comes to social listening [9,13,14,25], dialogue systems (chatbots), recommender systems, reputation monitoring, and the detection of fake news [10] and hate speech [8].

There do exist a few monolingual humour corpora exist: for example, the datasets created for shared tasks of the International Workshop on Semantic Evaluation (SemEval): #HashtagWars: Learning a Sense of Humor [23], Detection and Interpretation of English Puns [20], Assessing Humor in Edited News Headlines [12], and HaHackathon: Detecting and Rating Humor and Offense [17]. Mihalcea and Strapparava [18] collected 16,000 humorous sentences and an equal number of negative samples from news titles, proverbs, the British National Corpus, and the Open Mind Common Sense dataset, while another dataset contains 2,400 puns and non-puns from news sources, Yahoo! Answers, and proverbs [3,28]. Most datasets are in English, with some notable exceptions for Italian [24], Russian [1,6], and Spanish [2].

Strategies for Wordplay Translation. Humorous wordplay often exploits the confrontation of similar forms but different meanings, evoking incongruity between expected and presented stimuli, and this makes it particularly important in NLP to study the strategies that human translators use for dealing with it [4,27]. On the one hand, this is because MT is generally ignorant of pragmatics and assumes that words in the source text are formed and used in a conventional manner. MT systems fail to recognize the deliberate ambiguity of puns or the unorthodox morphology of neologisms, leaving such terms untranslated or else translating them in ways that lose the humorous aspect [19]. Apart from these implementation issues, human translation strategies could also inform the evaluation of machine-translated wordplay, since existing metrics based on lexical overlap [15,22] are not applicable.

Perhaps the most commonly cited typology of wordplay translation strategies is that of Dirk Delabastita [4,5]. This typology was developed on the basis of parallel corpus analysis and therefore reflects the techniques used by working translators. And while the typology was developed specifically for puns (a type of wordplay that exploits multiple meanings of a term or of similar-sounding words), many of the strategies are applicable to other forms not based on ambiguity. Delabastita’s basic options are the following:

- PUN→PUN:** The source text pun is translated by a target language pun.
- PUN→NON-PUN:** The pun is translated by a non-punning phrase, which may reproduce all senses of the wordplay or just one of them.
- PUN→RELATED RHETORICAL DEVICE:** The pun is replaced by some other rhetorically charged phrase (involving repetition, alliteration, rhyme, etc.)
- PUN→ZERO:** The part of text containing the pun is omitted altogether.
- PUN ST=PUN TT:** The pun is reproduced verbatim, without attempting a target-language rendering.
- NON-PUN→PUN:** A pun is introduced in the target text where no wordplay was present in the source text.
- ZERO→PUN:** New textual material involving wordplay is added in the target text, which bears no correspondence whatsoever in the source text.
- EDITORIAL TECHNIQUES:** Use of some paratextual strategy for explaining the pun of the source text (footnote, preface, etc.).

Delabastita insists on one further point: the techniques are by no means exclusive. A translator could, for instance, suppress a pun somewhere in their target text (PUN→NON-PUN), explain it in a footnote (EDITORIAL TECHNIQUES), then try to compensate for the loss by adding another pun somewhere else in the text (NON-PUN→PUN or ZERO→PUN). The very typology of translation strategies drawn in [5] directly points to the main reason for the difficulty of conceiving a working MT of puns. Translating wordplay does not involve recourse to what we may commonly think of as translation, but to (almost) autonomous creative writing activities starting from a situation determined by the source text. Therefore, a more realistic goal for NLP might probably be machine detection, followed by human or computer-assisted translation of wordplay.

3 Task setup

Data. Wordplay includes a wide variety of phenomena that exploit or subvert the phonological, orthographical, morphological, and semantic conventions of a language. We have collected over two thousand translated examples of wordplay, in English and French, from video games, literature, and other sources. Each example has been manually classified according to a well-defined, multi-label inventory of wordplay types and structures, and annotated according to its lexical-semantic or morphosemantic components.

The type inventory covers phenomena such as puns; alliteration, assonance, and consonance (repetition of sounds across nearby words); portmanteaux (combining parts of multiple words into a new word); malapropisms (the erroneous use

of a word in place of a similar-sounding one); spoonerisms (exchanging the initial sounds of nearby words); anagrams (a word or phrase formed by rearranging the letters of another); and onomatopoeia (a word coined to approximate some non-speech or non-language sound). The structure inventory can be used to further specify the entities involved in certain types of ambiguity-based wordplay; its labels include homophony (words with the same pronunciation but different spelling), homography (words with the same spelling but different pronunciation), homonymy (words with the same spelling and pronunciation), paronymy (words with different spelling and pronunciation), lexical structure (the figurative and literal readings of an idiom), morphological structure (different morphological analyses of the same word), and syntactic structure (when a word takes on different meanings according to how the wider phrasal context is parsed).

The examples in Table 1 give some idea of our annotated data, about half of which consists of proper nouns and neologisms.

Table 1. Examples of annotated instances of wordplay.

| Instance | Type | Structure | Interpretation |
|---|--|-----------|-----------------------|
| Why is music so painful? Because it HERTZ. | pun | paronymy | hurts/hertz |
| Weasleys' Wildfire Whiz-bangs | alliteration, assonance, and consonance | — | alliteration in 'w' |
| She was my secretariat. | malapropism | — | secretariat/secretary |

Evaluation Metrics. Pilot Task 1 includes both classification and interpretation components. Classification performance will be evaluated with respect to accuracy, while interpretation performance will be measured by exact-match comparison to the gold-standard annotations. Accuracy is preferable over precision, recall, and F_1 as the latter are designed for binary classification. For the same reason, they are not appropriate to evaluate translation quality.

For the wordplay translation tasks (Pilot Tasks 2 and 3), there do not yet exist any accepted metrics of translation quality. MT is traditionally measured with the BLEU (Bilingual Evaluation Understudy) metric, which calculates vocabulary overlap between the candidate translation and a reference translation [22]. However, this metric is clearly inappropriate for use with wordplay, where a wide variety of translation strategies (and solutions implementing those strategies) are permissible. And as our *Wonderland* example from §1 demonstrates, many of these strategies require metalexical awareness and preservation of features such as lexical ambiguity and phonetic similarity. (Consider how substituting the synonymous *leçons* for *cours* in Parisot's translation would lose the wordplay, and indeed render the translation nonsensical, yet still result in a near-perfect BLEU score with the original translation.) Furthermore, overlap measures operate only

on larger text spans and not on individual words, the morphological analysis of which can be crucial for neologisms.

Evaluation of human translation quality is similarly problematic, with past studies on wordplay translation (e.g., [16]) favouring qualitative rather than quantitative analyses, or else employing only subjective metrics such as “acceptability” or “successfulness”. Part of the goal of the JOKER workshop is to work towards the development of evaluation metrics for the automated translation of wordplay. To this end, human evaluators will manually annotate the submitted translations according to both subjective measures and according to more concrete features such as whether wordplay exists in the target text, whether it corresponds to the type used in the source text, whether the target text preserves the semantic field, etc. At the end of the workshop, we will look for correspondences between the concrete and subjective measures, and consider how the concrete measures that best correlate with subjective translation quality might be automated.

4 Conclusion

The JOKER project addresses the issue of European identity through the study of humour in a cross-cultural perspective. Its main objective is to study the strategies of localization of humour and wordplay and to create a multilingual parallel corpus annotated according to these strategies, as well as to rethink evaluation metrics. To this end, we are organizing the CLEF 2022 JOKER track, consisting of a workshop and associated pilot tasks on automatic wordplay analysis and translation. A further goal of the workshop is to unify the scientific community interested in automatic localization of humour and wordplay and to facilitate future work in this area. Our multilingual corpus will be made freely available to the research community (to the extent permitted by third-party copyrights), and this data and evaluation framework will be a step forward to MT models adapted for creative language.

Further details on the pilot tasks and on how to participate in the track can be found in the call for papers and guidelines on the JOKER website.¹¹ Please join this effort and contribute by working on our challenges!

Acknowledgements

This work has been funded in part by the National Research Agency under the program *Investissements d’avenir* (Reference ANR-19-GURE-0001) and by the Austrian Science Fund under project M2625-N31. JOKER is supported by *La Maison des sciences de l’homme en Bretagne*. We thank Adrien Couaillet and Ludivine Grégoire for data collection and the Master’s students in translation at Université de Bretagne Occidentale for maintaining the JOKER website.

¹¹ <http://www.joker-project.com/>

References

1. Blinov, V., Bolotova-Baranova, V., Braslavski, P.: Large Dataset and Language Model Fun-Tuning for Humor Recognition. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. pp. 4027–4032. Association for Computational Linguistics, Florence, Italy (2019). <https://doi.org/10.18653/v1/P19-1394>, <https://www.aclweb.org/anthology/P19-1394>
2. Castro, S., Chiruzzo, L., Rosá, A., Garat, D., Moncecchi, G.: A Crowd-Annotated Spanish Corpus for Humor Analysis. In: Proceedings of the Sixth International Workshop on Natural Language Processing for Social Media. pp. 7–11. Association for Computational Linguistics, Melbourne, Australia (Jul 2018). <https://doi.org/10.18653/v1/W18-3502>, <https://www.aclweb.org/anthology/W18-3502>
3. Cattle, A., Ma, X.: Recognizing Humour using Word Associations and Humour Anchor Extraction. In: Proceedings of the 27th International Conference on Computational Linguistics. pp. 1849–1858. Association for Computational Linguistics, Santa Fe, New Mexico, USA (Aug 2018), <https://www.aclweb.org/anthology/C18-1157>
4. Delabastita, D.: There’s a Double Tongue: An Investigation Into the Translation of Shakespeare’s Wordplay, with Special Reference to Hamlet. Rodopi, Amsterdam (1993)
5. Delabastita, D.: Wordplay as a translation problem: A linguistic perspective. In: 1. Teilband: Ein internationales Handbuch zur Übersetzungsforschung, pp. 600–606. De Gruyter Mouton (Jul 2008). <https://doi.org/10.1515/9783110137088.1.6.600>, <https://www.degruyter.com/document/doi/10.1515/9783110137088.1.6.600/html>
6. Ermilov, A., Murashkina, N., Goryacheva, V., Braslavski, P.: Stierlitz Meets SVM: Humor Detection in Russian. In: Ustalov, D., Filchenkov, A., Pivovarova, L., Žižka, J. (eds.) Artificial Intelligence and Natural Language. pp. 178–184. Communications in Computer and Information Science, Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-030-01204-5_17
7. Farwell, D., Helmreich, S.: Pragmatics-based MT and the Translation of Puns. In: Proceedings of the 11th Annual Conference of the European Association for Machine Translation. pp. 187–194 (Jun 2006), <http://www.mt-archive.info/EAMT-2006-Farwell.pdf>
8. Francesconi, C., Bosco, C., Poletto, F., Sanguinetti, M.: Error Analysis in a Hate Speech Detection Task: the Case of HaSpeeDe-TW at EVALITA 2018. In: Bernardi, R., Navigli, R., Semeraro*, G. (eds.) Proceedings of the 6th Italian Conference on Computational Linguistics (CLiC-it 2019) (Nov 2018), <http://ceur-ws.org/Vol-2481/paper32.pdf>
9. Ghanem, B., Karoui, J., Benamara, F., Moriceau, V., Rosso, P.: IDAT@FIRE2019: Overview of the Track on Irony Detection in Arabic Tweets. In: Mehta, P., Rosso, P., Majumder, P., Mitra, M. (eds.) Working Notes of FIRE 2019 – Forum for Information Retrieval Evaluation (Dec 2019)
10. Guibon, G., Ermakova, L., Seffih, H., Firsov, A., Noé-Bienvenu, G.L.: Multilingual Fake News Detection with Satire. In: Proceedings of the 20th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2019) (Apr 2019), <https://halshs.archives-ouvertes.fr/halshs-02391141>
11. Hong, B.A., Ong, E.: Automatically Extracting Word Relationships as Templates for Pun Generation. In: Proceedings of the Workshop on Computational Approaches to Linguistic Creativity. pp. 24–31. Association for Computational Linguistics, Boulder, Colorado (Jun 2009), <https://aclanthology.org/W09-2004>

12. Hossain, N., Krumm, J., Gamon, M., Kautz, H.: SemEval-2020 Task 7: Assessing Humor in Edited News Headlines. In: Proceedings of the Fourteenth Workshop on Semantic Evaluation. pp. 746–758. International Committee for Computational Linguistics (Dec 2020), <https://aclanthology.org/2020.semeval-1.98>
13. Karoui, J., Benamara, F., Moriceau, V., Patti, V., Bosco, C., Aussenac-Gilles, N.: Exploring the Impact of Pragmatic Phenomena on Irony Detection in Tweets: A Multilingual Corpus Study. In: 15th European Chapter of the Association for Computational Linguistics (EACL 2017). vol. 1 - long pap, pp. 262–272. Association for Computational Linguistics (ACL), Valencia, ES (2017), <https://oatao.univ-toulouse.fr/18921/>
14. Karoui, J., Farah, B., Moriceau, V., Aussenac-Gilles, N., Hadrich-Belguith, L.: Towards a Contextual Pragmatic Model to Detect Irony in Tweets. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers). pp. 644–650. Association for Computational Linguistics, Beijing, China (2015). <https://doi.org/10.3115/v1/P15-2106>, <http://aclweb.org/anthology/P15-2106>
15. Klakow, D., Peters, J.: Testing the correlation of word error rate and perplexity. *Speech Communication* **38**(1), 19–28 (Sep 2002). [https://doi.org/10.1016/S0167-6393\(01\)00041-3](https://doi.org/10.1016/S0167-6393(01)00041-3), <https://www.sciencedirect.com/science/article/pii/S0167639301000413>
16. Kolb, W., Miller, T.: Human-Computer Interaction in Pun Translation. In: Hadley, J., Taivalkoski-Shilov, K., Teixeira, C.S.C., Toral, A. (eds.) *Using Technologies for Creative-Text Translation*. Routledge (2022), to appear
17. Meaney, J.A., Wilson, S., Chiruzzo, L., Lopez, A., Magdy, W.: SemEval-2021 Task 7: HaHackathon, Detecting and Rating Humor and Offense. In: Proceedings of the 15th International Workshop on Semantic Evaluation. pp. 105–119. Association for Computational Linguistics (Aug 2021). <https://doi.org/10.18653/v1/2021.semeval-1.9>, <https://aclanthology.org/2021.semeval-1.9>
18. Mihalcea, R., Strapparava, C.: Making computers laugh: Investigations in automatic humor recognition. In: *Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing: Proceedings of the Conference*. pp. 531–538. Association for Computational Linguistics, Stroudsburg, PA (Oct 2005). <https://doi.org/10.3115/1220575.1220642>, <http://www.aclweb.org/anthology/H/H05/H05-1067>
19. Miller, T.: The Punster’s Amanuensis: The Proper Place of Humans and Machines in the Translation of Wordplay. In: Proceedings of the Second Workshop Human-Informed Translation and Interpreting Technology associated with RANLP 2019. pp. 57–65. Incoma Ltd., Shoumen, Bulgaria (Oct 2019). https://doi.org/10.26615/issn.2683-0078.2019_007, https://aclbg.org/proceedings/2019/RANLP_W1%202019/pdf/HiT-IT2019007.pdf
20. Miller, T., Hempelmann, C., Gurevych, I.: SemEval-2017 Task 7: Detection and Interpretation of English Puns. In: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017). pp. 58–68. Association for Computational Linguistics, Vancouver, Canada (2017). <https://doi.org/10.18653/v1/S17-2005>, <http://aclweb.org/anthology/S17-2005>
21. Nijholt, A., Niculescu, A., Valitutti, A., Banchs, R.E.: Humor in Human-Computer Interaction: A Short Survey. In: INTERACT 2017 (2017)
22. 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 as-

- sociation for computational linguistics. pp. 311–318. Association for Computational Linguistics (2002)
23. Potash, P., Romanov, A., Rumshisky, A.: SemEval-2017 Task 6: #HashtagWars: Learning a Sense of Humor. In: Proceedings of the 11th International Workshop on Semantic Evaluation. pp. 49–57 (Aug 2017). <https://doi.org/10.18653/v1/S17-2004>, <http://www.aclweb.org/anthology/S17-2004>
 24. Reyes, A., Buscaldi, D., Rosso, P.: An Analysis of the Impact of Ambiguity on Automatic Humour Recognition. In: Matoušek, V., Mautner, P. (eds.) Text, Speech and Dialogue. pp. 162–169. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04208-9_25
 25. Reyes, A., Rosso, P., Buscaldi, D.: From humor recognition to irony detection: The figurative language of social media. *Data & Knowledge Engineering* **74**, 1–12 (Apr 2012). <https://doi.org/10.1016/j.datak.2012.02.005>, <https://doi.org/10.1016/j.datak.2012.02.005>
 26. Valitutti, A., Toivonen, H., Doucet, A., Toivanen, J.M.: “Let Everything Turn Well in Your Wife”: Generation of Adult Humor Using Lexical Constraints. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). pp. 243–248. Association for Computational Linguistics, Sofia, Bulgaria (Aug 2013), <https://aclanthology.org/P13-2044>
 27. Vrticka, P., Black, J.M., Reiss, A.L.: The neural basis of humour processing. *Nature Reviews Neuroscience* **14**(12), 860–868 (Dec 2013). <https://doi.org/10.1038/nrn3566>, <https://www.nature.com/articles/nrn3566>
 28. Yang, D., Lavie, A., Dyer, C., Hovy, E.: Humor Recognition and Humor Anchor Extraction. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. pp. 2367–2376. Association for Computational Linguistics, Lisbon, Portugal (Sep 2015). <https://doi.org/10.18653/v1/D15-1284>, <https://www.aclweb.org/anthology/D15-1284>
 29. Yu, Z., Tan, J., Wan, X.: A Neural Approach to Pun Generation. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 1650–1660. Association for Computational Linguistics, Melbourne, Australia (Jul 2018). <https://doi.org/10.18653/v1/P18-1153>, <https://aclanthology.org/P18-1153>