Towards Generation and Recognition of Humorous Texts in Portuguese

Marcio Lima Inácio and **Hugo Gonçalo Oliveira**
CISUC - Centre for Informatics and Systems of the University of Coimbra
Department of Informatics Engineerings
Polo II, Pinhal de Marrocos, 3030-290
Coimbra, Portugal
{mlinacio, hroliv}@dei.uc.pt

Abstract

Dealing with humor is an important step to develop Natural Language Processing tools capable of handling sophisticated semantic and pragmatic knowledge. In this context, this PhD thesis focuses on the automatic generation and recognition of verbal punning humor in Portuguese, which is still an underdeveloped language when compared to English. One of the main goals of this research is to conciliate Natural Language Generation computational models with existing theories of humor from the Humanities while avoiding mere generation by including contextual information into the generation process. Another point that is of utmost importance is the inclusion of the listener as an active part in the process of understanding and creating humor; we hope to achieve this by using concepts from Recommender Systems in our methods. Ultimately, we want to not only advance the current state-of-the-art in humor generation and recognition, but also to help the general Portuguese-speaking research community with methods, tools and resources that may aid in the development of further techniques for this language. We also expect our systems to provide insightful ideas about how humor is created and perceived by both humans and machines.

1 Introduction

Natural Language Processing (NLP) research tends toward even more complex types of linguistic phenomena, requiring systems that are capable to deal with sophisticated semantic and pragmatic information (Cambria and White, 2014). To achieve such goals, it is essential to handle figurative and creative language (Reyes et al., 2012), which humor is a part of. Additionally, introducing the ability to recognize and create humor benefits general natural language-based systems, e.g. virtual agents, which can exploit such knowledge to make interaction with the user more pleasant and human-like, or news aggregators capable of dealing with satirical and comical articles.

This PhD thesis is also related to Computational Creativity (CC), a multidisciplinary field of research that is concerned with replicating, understanding and enhancing human creativity through computational tools (Veale and Pérez y Pérez, 2020).[1] In language, this has to deal with artifacts such as poetry, literature, and, especially for us, verbal humor. We refer to this intersection between NLP and CC with a focus on humor as Computational Humor Processing.

Although research on Computational Humor Processing dates back to the 1990s (Binsted and Ritchie, 1994; Ritchie, 1999), there is still much to advance on this area due to the complexity of the tasks involved. As mentioned by Amin and Burghardt (2020), the most successful systems for humor generation are based on predefined templates and rules, which is limiting in terms of linguistic realization. Furthermore, Clemêncio (2019) mention that humor recognition can be improved by exploiting new features and models, as well as through the creation of larger corpora with humorous texts, especially for the Portuguese language, the main focus of this proposal.

Within this context, this thesis proposes to tackle two main tasks related to computational humor processing: automatic recognition and generation of verbal humor, with a special focus on Portuguese, which still receives less attention than other more researched languages, such as English or Mandarin Chinese (Bender, 2019).

As verbal humor is a largely diverse phenomenon conveyed through many different ways, we decided to concentrate this research on a specific kind of humorous format: puns. This individ-

[1] For a more complete definition of Computational Creativity, we recommend the Association for Computational Creativity website: http://computationalcreativity.net/home/about/computational-creativity/. Accessed on: 28 nov. 2022.

Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop, pages 26–36

ual type of wordplay has had a special attention of the community as it is considered to be a simpler instance of linguistic artifact capable of expressing funniness exploiting word ambiguity (Kao et al., 2016). Nonetheless, this joke format still requires dealing with complex language information (e.g. common knowledge, semantic relations, linguistic realization, and surprise) which could be useful not only in Humor Processing, for instance systems that handle sarcasm, but also in other areas of NLP, such as Natural Language Understanding or Sentiment Analysis.

To this extent, our main research objectives are (i) to conciliate explicit theories from the Humanities with sophisticated computational models to exploit their different advantages to the tasks in matter, (ii) to avoid mere generation by incorporating contextual information to the creation of humorous puns, and (iii) to include the role of the user into the generation and understanding of humor by modelling their sense of humor and creating a personalized experience. As expected, the developed methods will be evaluated in comparison to existing techniques from the literature, which will be adapted to the Portuguese language if needed.

After this introduction, the remainder of the paper is organized as follows: in section 2 we present in more detail the concepts, motivations, and general background for this project, followed by a discussion about punning humor in section 3. The research proposal alongside the intended methodologies are mentioned in section 4. Finally, we include considerations about the limitations of the project and some ethical concerns that may rise from this research.

2 Background

This project has two main fronts, humor recognition and humor generation, which are discussed in more detail below.

2.1 Humor generation

Works on the computational processing of humor date back to 1994, with two prominent systems named LIBJOG (Raskin and Attardo, 1994) and JAPE (Binsted and Ritchie, 1994) created specifically for the creation of jokes based on explicit templates to be filled according to a predefined set of rules; this kind of rule-based approach has been the most used for the task (Ritchie et al., 2006; Stock and Strapparava, 2003; Winters et al., 2018). Later, Hong and Ong (2009) created T-PEG, a system capable of learning templates and rule sets automatically from given punning riddles.

Another approach explored for humor generation is through lexical replacement on input texts, for instance the methods by Valitutti et al. (2016) and He et al. (2019). Lastly the most recent systems use some modern Natural Language Generation (NLG) techniques to create punning jokes, for example via Neural Networks (Yu et al., 2018).

In their recent survey about humor generation methods, Amin and Burghardt (2020) point out that neural-based systems are considered by users to create texts with a higher degree of linguistic complexity, using a larger variety of textual devices to convey the intended message; however, they also state that such techniques are still far from achieving a decent level of humorousness when compared to the usual template-based approaches. These observations reveal that both techniques have their strengths and their shortcomings, opening some path for future research on the matter.

Some authors also advocate that typical NLG methods are not suitable for creating creative texts, as they are built to approximate the general patterns of language which is usually the opposite goal of linguistic creativity (He et al., 2019). Hempelmann (2008) even mentioned that the only way to capture complex irregular phenomena of language is through explicit linguistic theories besides Machine Learning (ML) algorithms. However, even though they might not be completely fit for complex semantic reasoning (Bender and Koller, 2020) or creative tasks, modern techniques, such as Large Language Models (LLMs), still show impressive results producing human-like texts (Stevenson et al., 2022), justifying some research on integrating both points of view for the proposed task.

Investigation about automatic humor generation can also be done toward the personalization of the content being produced. As mentioned by Siekiera et al. (2022), "the success of a joke is strongly cultural-based", which raises questions about the role of the listener in the process of creating and perceiving humor (Veale, 2004); from this point of view, the authors indicate that future work can take into account the cultural background — and we might also include personal preferences — of the user in computational tools for humor processing. This perspective is also shared by Winters et al. (2018), who mention that creating adaptive systems

capable of generating jokes based on the user's personal preferences is prone to outperform other methods that do not have such ability.

A further argument for the insertion of the user's preferences into humor generation process lies in the evaluation of such approaches. It is not unusual that authors mention that their proposed scores do not correlate with human ratings of funniness (Kao et al., 2016; He et al., 2019; Gonçalo Oliveira and Rodrigues, 2018). This may be due to the fact that human evaluation is usually done without taking into account the evaluators' personal preferences, resulting in a general neutral sentiment; for example, even for human-made jokes, He et al. (2019) report that the average rating obtained was 3 in a scale from 1 to 5. These observations show that including the reader actively in the process may be fruitful not only for the development of better methods, but also for the formulation of more robust evaluation methodologies for such systems.

2.2 Humor recognition

The recognition of humorous texts has been traditionally tackled as a binary classification problem dating back to the early 2000s (Yokogawa, 2002; Taylor and Mazlack, 2004; Mihalcea and Strapparava, 2005). Usually, such systems differ on their choices concerning the feature set used, as the ML algorithms are frequently the same: Naïve Bayes, Support Vector Machines (SVM), Decision Trees, and Random Forest. For example, Mihalcea and Strapparava (2005) features were primarily stylistic ones (alliteration, adult slangs, and antonymy) while Mihalcea and Pulman (2007) focus on semantic characteristics such as negations, negative human trains, and words related to professional communities. In their turn, Sjöbergh and Araki (2007) aim at using shallow textual features without any intent to capture meaning, identifying some relevant traits as frequent words, text similarity with known jokes, and idiomatic expressions.

Despite being the most common approach, some authors do not use ML for the task. An example is the work by Tinholt and Nijholt (2007) who create a rule-based system that recognizes humor potential in non-humorous texts by identifying ambiguous anaphora cases through a semantic graph. Alternately, Kao et al. (2016) developed probabilistic metrics for ambiguity and distinctiveness to recognize humor in homophonic and paronymic puns; their results were promising as the scores were successful in differentiating puns from non-puns, but determining the level of funniness of a joke is still a challenging task.

Another interesting observation by Kao et al. (2016) was that their measures gave hints about which words in the text were mostly related to the humorous effect, opening a new path of research toward not only identifying humor but also explaining why a certain text might be considered funny. This kind of task has also been approached by Yang et al. (2015), which use a funniness model to determine humor-inducing words.

As mentioned in subsection 2.1, humor processing systems usually do not take into account the user's specific preferences when creating jokes. For humor recognition, the same concepts may also be applied, it might be an interesting approach to not only identify whether an artifact is funny, but also to whom (or to which groups of people) it may have the intended humorous effect, in a manner similar to demographic filtering present in Recommender Systems (Bobadilla et al., 2013).

2.3 Computational humor processing in Portuguese

The majority of works in Computational Humor are focused on the English language, requiring large robust lexical resources — e.g. WordNet (Fellbaum, 1998) and ConceptNet (Speer et al., 2017) — and annotated corpora, which are not always available or fully developed for other languages such as Portuguese. Although there are initiatives to create the needed resources, they are still limited when compared to their English counterparts, which is a natural consequence of the smaller number of researchers interested in the Portuguese language.

For such reasons, Computational Humor systems for Portuguese are clearly in an early stage, with most generation approaches based on handcrafted templates and rules (Gonçalo Oliveira et al., 2016; Gonçalo Oliveira and Rodrigues, 2018). Alternately, Mendes and Gonçalo Oliveira (2020) propose a method to create humor by editing an input text. Humor recognition, in its turn, has been tackled through traditional ML with stylistic and semantic features (Clemêncio, 2019; Gonçalo Oliveira et al., 2020), similar to the previous works for the English language.

Data availability is also a concern, as there are few annotated corpora with humorous texts in Portuguese. For example, Gonçalo Oliveira et al.

(2020) introduced a collection of one-liners (short jokes) and riddles with a binary annotation created automatically according to their source. On a more broad relation to this project, some other corpora deal with other types of figurative language: Wick-Pedro et al. (2020) provide a corpus of tweets related to satirical news with a manual annotation on the intents of the users, as well as their sentiment toward the subject; furthermore, some authors also provide collections of user-generated content with annotations and linguistic descriptions about irony in such texts (Carvalho et al., 2009; de Freitas et al., 2014; Wick-Pedro et al., 2020).

In sum, working with a language as Portuguese is challenging due to the lack of large corpora, robust resources, and more modern methods to start with. Therefore, we believe that this thesis will be of great value, not only to research on Computational Humor Processing or Computational Creativity, but also to the general Portuguese-speaking NLP community.

3 Punning Humor

As mentioned by Kao et al. (2016), puns are a simpler instance of verbal humor based on phrase and word ambiguity, which makes them an ideal starting point for research on the area. This thought is shared by Aleksandrova (2022), who also mentions that puns are relevant due to their frequency in everyday live.

Even though they may be simpler and shorter, Hempelmann (2008) argues that punning jokes still require sophisticated models for doing meaningful research, as they contain all necessary elements to create a humorous effect, a sufficiently complex phenomenon by itself. Therefore, research on this kind of humor can produce knowledge that might be generalizable to other types of humor or serve as a basis for investigating other humor-related phenomena.

For those reasons, this project focuses exclusively on puns and punning jokes. Our working definition for this kind of verbal humor is as follows:

> *A pun is a form of wordplay in which one sign (e.g., a word or phrase) suggests two or more meanings by exploiting polysemy, homonymy, or phonological similarity to another sign, for an intended humorous or rhetorical effect.*
>
> Miller et al. (2017)

From the definition, punning humor is created through a relation between form (spoken or written) and meaning, requiring that a sign must evoke multiple meanings in the given context. Some examples of puns from Miller and Gurevych (2015) with different characteristics are presented below; the punning word is highlighted in bold and the specific relation is between parentheses.

1. A lumberjack's world revolves on its **axes**. (homography)
2. She fell through the window but felt no **pane**. (homophony)
3. A political prisoner is one who stands behind her **convictions**. (homonymy)
4. The sign at the nudist camp read, "**Clothed** until April." (paronymy)

It is important to stress that the pun relates to not only words but linguistic signs in general: word segments, phrases, acronyms, graphemes, onomatopoeias, and others, as illustrated by some of the examples provided by Aleksandrova (2022).

1. In English, we '**drive** cars on parkways' and 'park cars on **drive**ways'. (word segments)
2. What four letters frighten a thief? **O.I.C.U.** (phrase and graphemes)
3. How much space will Brexit free up in the European Union? 1 **GB**. (acronym)

For the reasons mentioned, in this thesis, we will focus on puns and punning humor. With this, we hope to advance the current research on the computational processing of verbal humor in Portuguese, as we will elaborate further in the next sections.

4 Research Proposal

The main objective of this thesis is to **develop methods and resources for the computational recognition, analysis, and generation of verbal punning humor in Portuguese**. To this extent, we defined some specific goals to be reached throughout the development of the research work.

- Develop, evaluate and, if needed, adapt to the Portuguese language existing methods for the automatic recognition and generation of puns;

- Create a corpus of short punning jokes in Portuguese with user ratings on their funniness alongside annotations of humor-inducing words;
- Create new methods for pun generation and recognition by determining and adapting linguistic and psycholinguistic theories of humor, surprise, or creativity to a computational scenario, combining them with different approaches for NLG and ML, especially LLMs;
- Avoid mere generation by including contextual information — automatically generated or not — to novel or existing methods for computational humor generation systems;
- Include specificities of the target audience in the process of generation, evaluation, and ranking of punning humor through concepts of Recommender Systems filtering;
- Evaluate the proposed techniques, comparing them against each other, existing methods, and baseline systems.

To come to each objective, we defined some methodologies to be followed during the research, which will be discussed as follows.

4.1 Adaptation of the literature methods

Since most of the techniques for computational humor processing are based on rules or large robust resources, they tend to be limited to a single language, usually English. To provide a fair comparison with our novel methods, and also to stimulate research on rules and resources for the Portuguese language, our first objective is to select, implement and adapt systems from the literature to our working language.

For pun recognition, as mentioned in subsection 2.3, most of this work for the ML-based methods has been done by Gonçalo Oliveira et al. (2020); however, we might still implement other methods, such as the ones by Yang et al. (2015) and Kao et al. (2016), which also incorporate to some extent the tasks of automatically determining levels of funniness and identifying humor triggering words in the input text.

For the creation of puns, there is still much work to be done. The first method that seems interesting to be adapted is the one by Hong and Ong (2009), which automatically learns templates and rules from pre-existing puns. This system is especially challenging, as it relies on some specific resources for phonological, lexical, and semantic analyses that should have their counterparts in Portuguese. Other techniques are the ones by Yu et al. (2018), a recurrent neural network to create homographic puns, and He et al. (2019), which is based on probabilistic models of surprise to edit input texts to create punning humor.

Implementing such systems for a language other than English will help to start advancing the current studies on the matter and also bring attention to specificities of the Portuguese language and its resources that need to be taken into account when developing further methods.

4.2 Corpus Creation

A key point of this thesis will be the creation of a corpus of short punning jokes in Portuguese, which will enable not only our research on automatic pun identification and generation, but also linguistic studies on this format of verbal humor. We intend to make the corpus publicly available, alongside every annotation that results from this project.

The corpus will be gathered manually from websites, social media, and YouTube videos, following some guidelines regarding the definition of punning humor by Miller et al. (2017) (section 3) and the textual format aimed for: short texts capable of being written in a single line, this means that dialogues or narrative arcs will not be included in this first version of the corpus. We will also provide a classification following the taxonomy defined by Hempelmann and Miller (2017), explicitly marking homophony and homography, which, despite not being our main focus, might help other researchers to better filter the data for their analyses.

During the data gathering, there will surely be some hard cases, i.e. texts in which the gatherer has some doubt about the nature of the humorous effect or if the instance should be included into the corpus, needing to refer back to the guidelines document. Such cases will be highlighted to enable a deeper discussion about what is punning humor, how it occurs in general, and how we created our corpus. As this is an ongoing work, we have already found some interesting examples, presented below in Table 1, that will need to be further analyzed to be discussed in deeper detail.

Table 1: Example of hard cases from the ongoing corpus collection, including onomatopoeias, neologisms, foreign languages, and others.

Original joke in Portuguese	English translation	Comments
Qual é a consola de jogos preferida dos polícias? Wii U! Wii U! Wii U! Wii U!	*What is the policemen favorite video game console? Wii U! Wii U! Wii U! Wii U!*	This joke uses an onomatopoeia, as the sound of "Wii U" resembles the sound of the sirens used in police cars.
Que nome se dá a uma freira no casino? Católica apostólica.	*How does one call a nun in a casino? Catholic Apostolic.*	This joke turns an existing word, "apostólica" (*apostolic*), into a neologism that relates to the concepts of "aposta" (*bet*) and "alcoólica" (*alcoholic*), creating a new word that describes a person addicted to gambling.
Como é que se diz "fim" em japonês? Sakabô.	*How does one say "end" in Japanese? Sakabô.*	This pun creates the word "sakabô", whose sounds resemble a foreign language (Japanese) and approximate the pronunciation of "se acabou" (*it is over*).
O que diz um castor excitado? Suck my dique!	*What does a horny beaver say? Suck my dam.*	This text uses a mix of languages, Portuguese and English, taking advantage of the similairy in pronunciation of the words "dique" (*dam*) in Portuguese and *dick* in English.
Sonhei que pesava menos de uma milésima de grama. E fiquei tipo "0mg".	*I had a dream that I wheighted less than one milligram. And I was like "0mg".*	This joke uses the written resemblance of not entire words, but the graphical symbols themselves: 0 (*zero*) and O (the letter *O*), to create a pun between "0mg" (*zero milligrams*) and "omg" (acronym for *oh my god*).

4.3 Use of explicit theories

In this project, we share the points of view by Hempelmann (2008) and Amin and Burghardt (2020) that explicit linguistic theories have much to offer when dealing with tasks that handle complex irregular phenomena of the language, such as creativity and humor. Nonetheless, as noted by Stevenson et al. (2022), the power of LLMs to create human-like linguistically complex text is too strong even for creative tasks, but with clear limitations on how creative or funny their outputs are. This indicates that it might be fruitful to use such explicit theories to overcome these limitations, for example by including richer knowledge in prompts or by using Augmented Language Model (ALM) techniques (Mialon et al., 2023).

For humor detection, Kao et al. (2016) show that linguistic-inspired scores can be a promising path for research, capable not only of differentiating puns from non-puns, but also to give hints about which are the words in the text mostly related to the proposed humorous effect. Such results can be combined with features from the literature known to be effective in the task — such as ambiguity and taboo language — as well as with novel latent semantic and language models that have been achieving impressive results in various NLP tasks (Bender and Koller, 2020).

Accordingly, for our second objective, we de-

cided to create new methods for the computational processing of puns by exploiting concepts and findings of explicit theories from the Humanities (Linguitics, Psycholinguistics, Cognitive Linguistics, Psychology, and others) to help overcome the limitations of existing computational models for the processing of language.

Some examples of theories that can be studied and exploited in our research are: Script-based Semantic Theory of Humor (SSTH; Raskin, 1984), General Theory of Verbal Humor (GTVH; Attardo and Raskin, 1991), Optimal Innovation Hypothesis (Giora et al., 2004), models for surprise (Macedo and Cardoso, 2001; Tobin, 2018; Chieppe et al., 2022), and theories on sense of humor (Martin, 2003).

4.4 Avoiding mere generation

The majority of systems that generate humorous texts do not take into account some contextual information to constrain the creation process, a phenomenon that we call "mere generation" (Ventura, 2016). For example, the method developed by Winters and Delobelle (2021) uses Language Models and Genetic Algorithms to edit news headlines to make them sound funnier, however there is no effort to use the actual text of the article to ensure that the title will match to its content, which is a desirable characteristic in a real-life scenario.

Another instance, for the Portuguese language, is the SECO system (Gonçalo Oliveira and Rodrigues, 2018) that creates funny riddles from a list of compound words through pre-defined templates and rules. The tool exhaustively tries to create a joke for every entry in the list and each template, regardless of any input about the content or topic of the intended output. This process generates a large amount of riddles that are not necessarily funny nor suitable for any unrestricted context.

To make our methods more fit to final applications, one of the objectives of this thesis is to avoid mere generation by including contextual information to constrain the generation process. This can be achieved through keywords, topics, conversation utterances, narrative texts, news articles, and so on.

4.5 Include the user in the process

As mentioned by Winters et al. (2018), "*an integrated humor generator that is capable of generating jokes adapted to the user might [...] outperform a generator that does not possess this capability.*" This quote explicits the necessity of considering the sense of humor of the listener — the user — in Computational Humor Processing systems. Something along these lines was also discussed by Veale (2004) when the author mentions that an essentialist view of humor, i.e. an interpretation of humor simply as a result of language-related characteristics, is insufficient to deal with the complexity of this phenomenon.

In this context, one of the main objectives of this thesis is to create methods for both pun recognition and pun generation that take into account the user's personal sense of humor to create a more personalized experience. As the two tasks are distinct, we give more details on them separately.

Pun recognition Few jokes are considered universally funny, as their humorous effect is dependent not only on their textual characteristics, but also on the cultural and personal background of the listener; therefore, systems that categorically determine if some artifact is funny or not might be flawed (or at least limited). To deal with this, we might create systems that, besides predicting if a text is funny, also detect which users or demographic groups might perceive them as so.

Pun generation For our second task, some model of the user's interests should constrain, prime, or guide the generation process to create puns with a higher chance of producing a funny outcome for that specific person.

In both cases, we need to capture the user's preferences into a model to be used by the systems in their specific tasks. For this modelling, Recommender Systems (RS) seem to be an interesting field of research that has much to offer in terms of concepts, techniques, and scores (Bobadilla et al., 2013).

As in this research we deal with an indefinite set of artifacts, using content filtering is essential; this is a category of methods for RS that focus on recommending items according to similarities between their content (text, image, sound, etc.) and the user's profile created from their previous choices (purchases, likes, shares, and others). Additionally, demographic filtering can also be valuable to identify social groups prone to find some artifact funny by analyzing personal characteristics of individuals, such as age, gender, country, and the like.

There exist other filtering techniques — e.g. col-

laborative and social filtering — however, we believe that the ones mentioned above are a satisfying starting point to include into our methods for Computational Humor Processing so that the user has a more active role in such processes.

More details about how these techniques are to be used depend on the amount of human resources we will have available, as datasets on RS usually deal with thousands of users. For example, Jester (Goldberg et al., 2001), a well-known data set for joke recommendation, has 100 jokes evaluated by 73,421 users in their original version.

Another possibility to achieve this goal is to use Reinforcement Learning techniques, in which the model is fine-tuned according to a user feedback (ranking, scoring, or classification) to result in a modified system that takes into account their personal tastes.

4.6 Evaluation

The evaluation of humor processing systems is extremely complex, the whole phenomenon alike; therefore, automatic scores are scarce, especially for generation. As our two main tasks require different evaluation process, we detail them separately below.

Pun recognition As humor recognition is usually interpreted as a classification or regression task, all available automatic evaluation methods can be used, such as precision, recall, accuracy, Mean Square Error (MSE), Mean Absolute Error (MAE), R Square (R^2), and so on. However, when including the user as an active part of the process (according to our intentions stated in subsection 4.5), the evaluation becomes more complex and needs to be further thought of carefully.

Pun generation As reported by Amin and Burghardt (2020), the evaluation of humor generation systems does not count with any satisfying widely-adopted automatic scoring method; therefore, this process is in general performed manually. A possibility is to evaluate the generated texts according to the five criteria used by Gonçalo Oliveira and Rodrigues (2018): interpretation, surprise, novelty, and humor. It is also possible to analyze the linguistic complexity of the outputs, as previously done by Amin and Burghardt (2020). Since the evaluation will be made mainly by hand, the quality of our user-focused pun generation methods will naturally take into account their personal point of view and interests, meaning that this evaluation procedure seems suitable to every system we might create for this task. On the other hand, we may study the possibility of using the developed Pun Recognition systems to evaluate the generation results automatically.

In sum, in this research project, we will focus on already established evaluation methods used by the scientific community. However, there are cases in which we might need to develop custom methodologies to better attain valuable and fair evaluations of the proposed methods.

4.7 Expected results

As main results, we expect to develop computational methods for automatically generating, identifying, analyzing, and, possibly, evaluating punning humor in Portuguese. Such techniques can be incorporated into general applications to aid the final user in their needs to create and recognize humorous texts. Additionally, we believe that our systems may also help other researchers to better understand how verbal humor is created and perceived by both machines and humans.

Another goal we would like to achieve is to bring awareness to the Academia about the importance of multidisciplinarity in NLP and how research on Humanities may help to overcome the limitations of usual computational methods. Besides the already mentioned expectations, we ultimately hope that the resources and tools created during this research project help the community on a wide range of NLP problems, especially for the Portuguese language.

Limitations

As every research, this thesis has challenges and limitations regarding its execution and methodological decisions, which are discussed below.

The first limitation we point out is that this research focuses mainly on a very specific kind of verbal humor: puns; in addition, we aim only at short punning jokes. Therefore, we might not deal directly with a large amount of other phenomena, such as metaphors, narratives, dialogues, and so on. Nonetheless, as argued in section 3, this is a good starting point to advance the research on Computational Humor Processing for Portuguese.

Finally, the last limitation to which we call attention concerns the usage of concepts from Recommender Systems in the developed methods. There are some issues with RS, which our techniques will

probably be subject to, especially the cold-start problem, which occurs when it is not possible to provide reliable recommendations for new items or new users, due to a lack of initial ratings.

Ethics Statement

Despite having positive effects, such as promoting solidarity, bringing people together, and creating social acceptance and approval, humor can also be harmful, as it can be used as a form of social control, a correction for deviant behaviors, or as a way to legitimize social prejudice and stereotypes against marginalized groups (Crawford, 2003; Kuipers, 2008; Bemiller and Schneider, 2010).

Additionally, there are jokes which are not suitable for specific vulnerable groups, such as children, due to their possibly problematic content, e.g. sexual relations, pedophilia, harassment, xenophobic stereotypes, etc.

Therefore, it is important to bear these aspects of humor in mind throughout the whole research, including the data collection and the development of our methods. This will help to bring awareness and raise questions about how these systems, corpora, and resources might affect society not only in a positive light but also from a critical point of view.

Acknowledgments

This work is funded by national funds through the FCT – Foundation for Science and Technology, I.P. (grant number UI/BD/153496/2022), within the scope of the project CISUC (UID/CEC/00326/2020) and by the European Social Fund, through the Regional Operational Program Centro 2020.

References

Elena Aleksandrova. 2022. Pun-based jokes and linguistic creativity: Designing 3R-module. *The European Journal of Humour Research*, 10(1):88–107.

Miriam Amin and Manuel Burghardt. 2020. A survey on approaches to computational humor generation. In *Proceedings of the the 4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 29–41, Online. International Committee on Computational Linguistics.

Salvatore Attardo and Victor Raskin. 1991. Script theory revis(it)ed: Joke similarity and joke representation model. *Humor - International Journal of Humor Research*, 4(3-4).

Michelle L. Bemiller and Rachel Zimmer Schneider. 2010. IT'S NOT JUST A JOKE. *Sociological Spectrum*, 30(4):459–479.

Emily M. Bender. 2019. The #BenderRule: On Naming the Languages We Study and Why It Matters. The Gradient.

Emily M. Bender and Alexander Koller. 2020. Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5185–5198, Online. Association for Computational Linguistics.

Kim Binsted and Graeme Ritchie. 1994. An implemented model of punning riddles. In *Proceedings of the 12th National Conference on Artificial Intelligence*, volume 1, pages 633–638, Seattle. AAAI Press / The MIT Press.

J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez. 2013. Recommender systems survey. *Knowledge-Based Systems*, 46:109–132.

Erik Cambria and Bebo White. 2014. Jumping NLP Curves: A Review of Natural Language Processing Research. *IEEE Computational Intelligence Magazine*, 9(2):48–57.

Paula Carvalho, Luís Sarmento, Mário J. Silva, and Eugénio de Oliveira. 2009. Clues for detecting irony in user-generated contents: Oh...!! it's "so easy" ;-). In *Proceeding of the 1st International CIKM Workshop on Topic-sentiment Analysis for Mass Opinion - TSA '09*, page 53, Hong Kong, China. ACM Press.

Patrick Chieppe, Penny Sweetser, and Eryn Newman. 2022. Bayesian Modelling of the Well-Made Surprise. In *International Conference on Computational Creativity*, Bolzano-Bozen. Association for Computational Creativity (ACC).

André Clemêncio. 2019. *Reconhecimento Automático de Humor Verbal*. MSc, Universidade de Coimbra, Coimbra.

Mary Crawford. 2003. Gender and humor in social context. *Journal of Pragmatics*, 35(9):1413–1430.

Larissa A. de Freitas, Aline A. Vanin, Denise N. Hogetop, Marco N. Bochernitsan, and Renata Vieira. 2014. Pathways for irony detection in tweets. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing*, pages 628–633, Gyeongju Republic of Korea. ACM.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. Language, Speech, and Communication. MIT Press, Cambridge, Mass.

Rachel Giora, Ofer Fein, Ann Kronrod, Idit Elnatan, Noa Shuval, and Adi Zur. 2004. Weapons of Mass Distraction: Optimal Innovation and Pleasure Ratings. *Metaphor and Symbol*, 19(2):115–141.

Ken Goldberg, Theresa Roeder, Dhruv Gupta, and Chris Perkins. 2001. Eigentaste: A Constant Time Collaborative Filtering Algorithm. *Information Retrieval*, 4(2):133–151.

Hugo Gonçalo Oliveira, André Clemêncio, and Ana Alves. 2020. Corpora and baselines for humour recognition in Portuguese. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 1278–1285, Marseille, France. European Language Resources Association.

Hugo Gonçalo Oliveira, Diogo Costa, and Alexandre Miguel Pinto. 2016. One does not simply produce funny memes! - Explorations on the Automatic Generation of Internet humor. In *Proceedings of the Seventh International Conference on Computational Creativity*, pages 238–245, Paris. Sony CSL Paris, France.

Hugo Gonçalo Oliveira and Ricardo Rodrigues. 2018. Explorando a Geração Automática de Adivinhas em Português. *Linguamática*, 10(1):3–18.

He He, Nanyun Peng, and Percy Liang. 2019. Pun Generation with Surprise. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pages 1734–1744, Minneapolis. Association for Computational Linguistics.

Christian F. Hempelmann. 2008. Computational humor: Beyond the pun? In *The Primer of Humor Research*, number 8 in Humor Research, pages 333–360. Victor Raskin, Berlin, New York.

Christian F. Hempelmann and Tristan Miller. 2017. Puns: Taxonomy and Phonology. In *The Routledge Handbook of Language and Humor*, first edition, Routledge Handbooks in Linguistics, pages 95–108. Routledge.

Bryan Anthony Hong and Ethel Ong. 2009. Automatically extracting word relationships as templates for pun generation. In *Proceedings of the Workshop on Computational Approaches to Linguistic Creativity*, pages 24–31, Boulder, Colorado. Association for Computational Linguistics.

Justine T. Kao, Roger Levy, and Noah D. Goodman. 2016. A Computational Model of Linguistic Humor in Puns. *Cognitive Science*, 40(5):1270–1285.

Giselinde Kuipers. 2008. The Sociology of Humor. In *The Primer of Humor Research*, number 8 in Humor Research, pages 361–398. Victor Raskin, Berlin, New York.

Luís Macedo and Amílcar Cardoso. 2001. Modeling Forms of Surprise in an Artificial Agent. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 23.

Rod Martin. 2003. Sense of humor. In Shane J. Lopez and C. R. Snyder, editors, *Positive Psychological Assessment: A Handbook of Models and Measures.*, pages 313–326. American Psychological Association, Washington.

Rui Mendes and Hugo Gonçalo Oliveira. 2020. TECo: Exploring word embeddings for text adaptation to a given context. In *Proceedings of the Eleventh International Conference on Computational Creativity*, pages 185–188, Coimbra. Association for Computational Creativity (ACC).

Grégoire Mialon, Roberto Dessì, Maria Lomeli, Christoforos Nalmpantis, Ram Pasunuru, Roberta Raileanu, Baptiste Rozière, Timo Schick, Jane Dwivedi-Yu, Asli Celikyilmaz, Edouard Grave, Yann LeCun, and Thomas Scialom. 2023. Augmented Language Models: A Survey. *arXiv*.

Rada Mihalcea and Stephen Pulman. 2007. Characterizing Humour: An Exploration of Features in Humorous Texts. In Alexander Gelbukh, editor, *Computational Linguistics and Intelligent Text Processing*, volume 4394, pages 337–347. Springer Berlin Heidelberg, Berlin, Heidelberg.

Rada Mihalcea and Carlo Strapparava. 2005. Making computers laugh: Investigations in automatic humor recognition. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 531–538, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

Tristan Miller and Iryna Gurevych. 2015. Automatic disambiguation of English puns. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 719–729, Beijing, China. Association for Computational Linguistics.

Tristan Miller, Christian Hempelmann, and Iryna Gurevych. 2017. SemEval-2017 Task 7: Detection and Interpretation of English Puns. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pages 58–68, Vancouver, Canada. Association for Computational Linguistics.

Jonathan D. Raskin and Salvatore Attardo. 1994. Non-literalness and non-bona-fide in language: An approach to formal and computational treatments of humor. *Pragmatics & Cognition*, 2(1):31–69.

Victor Raskin. 1984. *Semantic Mechanisms of Humor*. Springer Netherlands, Dordrecht.

Antonio Reyes, Paolo Rosso, and Davide Buscaldi. 2012. From humor recognition to irony detection: The figurative language of social media. *Data & Knowledge Engineering*, 74:1–12.

Graeme Ritchie. 1999. Developing the Incongruity-Resolution Theory. Technical Report EDI-INF-RR-0007, Division of Informatics, University of Edinburgh, Edinburgh.

Graeme D Ritchie, Ruli Manurung, Helen Pain, Annalu Waller, and Dave O'Mara. 2006. The STANDUP interactive riddle builder. *IEEE Intelligent Systems*, 21(2):67–69.

Julia Siekiera, Marius Köppel, Edwin Simpson, Kevin Stowe, Iryna Gurevych, and Stefan Kramer. 2022. Ranking Creative Language Characteristics in Small Data Scenarios. In *International Conference on Computational Creativity*, Bolzano-Bozen. Association for Computational Creativity (ACC).

Jonas Sjöbergh and Kenji Araki. 2007. Recognizing Humor Without Recognizing Meaning. In Francesco Masulli, Sushmita Mitra, and Gabriella Pasi, editors, *Applications of Fuzzy Sets Theory*, volume 4578, pages 469–476. Springer Berlin Heidelberg, Berlin, Heidelberg.

Robyn Speer, Joshua Chin, and Catherine Havasi. 2017. ConceptNet 5.5: An open multilingual graph of general knowledge. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence and the Twenty-ninth Innovative Applications of Artificial Intelligence Conference*, 4444-4451.

Claire Stevenson, Iris Smal, Raoul Grasman, Matthijs Baas, and Han Van Der Maas. 2022. Putting GPT-3's Creativity to the (Alternative Uses) Test. In *International Conference on Computational Creativity*, Bolzano-Bozen. Association for Computational Creativity (ACC).

Oliviero Stock and Carlo Strapparava. 2003. HAHAcronym: Humorous Agents for Humorous Acronyms. *Humor - International Journal of Humor Research*, 16(3).

J.M. Taylor and L.J. Mazlack. 2004. Humorous wordplay recognition. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*, pages 3306–3311, The Hague, Netherlands. IEEE.

Hans Wim Tinholt and Anton Nijholt. 2007. Computational Humour: Utilizing Cross-Reference Ambiguity for Conversational Jokes. In Francesco Masulli, Sushmita Mitra, and Gabriella Pasi, editors, *Applications of Fuzzy Sets Theory*, volume 4578, pages 477–483. Springer Berlin Heidelberg, Berlin, Heidelberg.

Vera Tobin. 2018. *Elements of Surprise: Our Mental Limits and the Satisfactions of Plot*. Harvard University Press, Cambridge, Massachusetts.

Alessandro Valitutti, Antoine Doucet, Jukka M. Toivanen, and Hannu Toivonen. 2016. Computational generation and dissection of lexical replacement humor. *Natural Language Engineering*, 22(5):727–749.

Tony Veale. 2004. Incongruity in humor: Root cause or epiphenomenon? *Humor - International Journal of Humor Research*, 17(4).

Tony Veale and Rafael Pérez y Pérez. 2020. Leaps and Bounds: An Introduction to the Field of Computational Creativity. *New Generation Computing*, 38(4):551–563.

Dan Ventura. 2016. Mere generation: Essential barometer or dated concept. In *Proceedings of the Seventh International Conference on Computational Creativity*, pages 17–24. Sony CSL, Paris.

Gabriela Wick-Pedro, Roney L. S. Santos, Oto A. Vale, Thiago A. S. Pardo, Kalina Bontcheva, and Carolina Scarton. 2020. Linguistic Analysis Model for Monitoring User Reaction on Satirical News for Brazilian Portuguese. In Paulo Quaresma, Renata Vieira, Sandra Aluísio, Helena Moniz, Fernando Batista, and Teresa Gonçalves, editors, *Computational Processing of the Portuguese Language*, volume 12037, pages 313–320. Springer International Publishing, Cham.

Thomas Winters and Pieter Delobelle. 2021. Survival of the Wittiest: Evolving Satire with Language Models. In *Proceedings of the Twelfth International Conference on Computational Creativity*, pages 82–86, Mexico City, Mexico. Association for Computational Creativity (ACC).

Thomas Winters, Vincent Nys, and Daniel De Schreye. 2018. Automatic Joke Generation: Learning Humor from Examples. In Norbert Streitz and Shin'ichi Konomi, editors, *Distributed, Ambient and Pervasive Interactions: Technologies and Contexts*, volume 10922, pages 360–377. Springer International Publishing, Cham.

Diyi Yang, Alon Lavie, Chris Dyer, and Eduard Hovy. 2015. Humor Recognition and Humor Anchor Extraction. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2367–2376, Lisbon, Portugal. Association for Computational Linguistics.

T. Yokogawa. 2002. Japanese pun analyzer using articulation similarities. In *2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No.02CH37291)*, volume 2, pages 1114–1119, Honolulu, HI, USA. IEEE.

Zhiwei Yu, Jiwei Tan, and Xiaojun Wan. 2018. A Neural Approach to Pun Generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1650–1660, Melbourne, Australia. Association for Computational Linguistics.