The Punster's Amanuensis:
The Proper Place of Humans and Machines in the Translation of Wordplay

Tristan Miller

Austrian Research Institute for Artificial Intelligence (OFAI)
Freyung 6, 1010 Vienna, Austria
`tristan.miller@ofai.at`

Abstract

The translation of wordplay is one of the most extensively researched problems in translation studies, but it has attracted little attention in the fields of natural language processing and machine translation. This is because today's language technologies treat anomalies and ambiguities in the input as things that must be resolved in favour of a single "correct" interpretation, rather than preserved and interpreted in their own right. But if computers cannot yet process such creative language on their own, can they at least provide specialized support to translation professionals? In this paper, I survey the state of the art relevant to computational processing of humorous wordplay and put forth a vision of how existing theories, resources, and technologies could be adapted and extended to support interactive, computer-assisted translation.

1 Introduction

The creative language of humour and wordplay is all around us: every day we are amused by clever advertising slogans for which companies have paid vast sums to copy writers; our televisions and cinemas play an endless string of comedies, most of which get dubbed or subtitled into many different languages; and literary critics and scholars write volumes cataloguing and analyzing the wit of contemporary and classic authors. The ubiquity of humour and wordplay, and the constant need for creative professionals to evaluate, analyze, and translate it, would seem to make it a prime candidate for natural language processing (NLP) techniques such as machine translation (MT).

But despite being a recurrent and expected feature of many discourse types, humour and wordplay are necessarily out of scope for most real-world NLP applications. This is because these applications can only rigidly apply a fixed set of hand-crafted or automatically learned rules about the vocabulary, grammar, and semantics of a language. While these approaches work well enough on conventional language, they cannot robustly deal with texts that deliberately disregard or subvert linguistic conventions for a rhetorical effect. To computers, anomalies and ambiguities in the input, if they are detected at all, are seen as something that must always be resolved in favour of a single "correct" interpretation, rather than preserved and interpreted in their own right. For example, to native English speakers it is clear that the bank slogan "We feel loanley" contains a play on the words *loan* and *lonely*, but MT systems are stymied by the nonce term, leaving it untranslated, or else wrongly assuming it is a misspelling of *lonely* and losing the double meaning.

Recent years have seen a small flurry of NLP research aimed at changing the way computers process language by allowing them to recognize and interpret intentionally humorous ambiguity. While this work has laid some important groundwork, it is clear that there can never be a fully automatic, "one-size-fits-all" approach. Each expert user, whether a copy writer, a translator, or a literary scholar, has their own tasks, workflows, strategies, and goals. Customizing existing NLP tools to expert tasks has traditionally taken the form of automatically adapting existing data models to new languages and domains, or learning new data models with the help of "user-in-the-loop" techniques such as reinforcement learning. But neither of these approaches works around the "rigid rule" problem mentioned above: wordplay is by definition unpredictable and irreverent of rules and norms, and so cannot be easily captured in a predictive model.

In this paper, I survey the state of the art in

linguistics, computational linguistics, translation, and machine translation as it relates to humour and wordplay. On the basis of these findings, I argue that the proper place of machines in the translation of humorous wordplay is to support rather than replace human translators. In the vein of the *Translator's Amanuensis* proposed by Kay (1980), I make some specific proposals concerning how the hitherto disparate work in these fields can be connected with a view to producing "machine-in-the-loop" tools to assist human translators in selecting and implementing appropriate translation strategies for instances of puns and other forms of wordplay.

2 State of the Art

2.1 Linguistic Conceptions of Humour and Punning

The linguistic mechanisms of verbal humour have been studied since antiquity, by which time the roles of ambiguity and incongruity had already been recognized (Attardo, 1994, Ch. 1). Modern linguistics has significantly broadened and deepened this understanding, giving rise to formal theories of humour—that is, testable explanations of the necessary and sufficient linguistic conditions for a text to be humorous. Perhaps the most widely accepted of these today are Raskin's (1985) Script-based Semantic Theory of Humour (SSTH) and its extension, the Generalized Theory of Verbal Humour (GTVH; Attardo and Raskin, 1991).

Both the SSTH and the GTVH are based on the notion of *scripts*, or *semantic frames* as they are more commonly called in computational linguistics. A script is a collection of semantic information, internalized by a native speaker, that specifies "characteristic features, attributes, and functions of a denotatum, and its characteristic interactions with things necessarily or typically associated with it" (Alan, 2001, p. 251). Under the SSTH, humour is evoked when a given text is compatible, at least in part, with two different scripts that are "opposite" in some culturally significant sense (e.g., life vs. death). The GTVH incorporates this notion of *script opposition* (SO) as the first of six parameters, or *knowledge resources* (KRs), that, when instantiated, uniquely characterize a given joke. The other five KRs are, in descending order of salience, the *logical mechanism* (LM), the (often faulty) reasoning whereby the incongruity of the scripts is resolved; the *situation* (SI), the non-humorous set-

Figure 1: A multimodal pun from the 1984 film *Top Secret!*: Hearing it cough, the woman asks the driver if the horse is alright. The driver replies, "Oh, he caught a cold the other day and he's just a little hoarse."

ting and paraphernalia of the joke; the *target* (TA) or butt of the joke; the *narrative strategy* (NS) or "genre" of the joke; and *language* (LA), the lexical, syntactic, phonological, and other linguistic choices that express the other KRs. Empirical validation has shown the postulated dependency hierarchy of the six KRs to be fundamentally correct (Ruch et al., 1993), which has important implications (discussed in §2.3) for the translation of humour.

Punning is a form of language play in which a word or phrase is used to evoke the meaning of another word or phrase with a similar or identical pronunciation. The term *pun* can refer to such an instance of wordplay as a whole, or more specifically to the word/phrase in it with the more salient meaning; the *target*[1] is the secondary word/phrase that is evoked. Figure 1 presents an example for the pun *hoarse* and the target *horse*.

Puns are one of the most studied phenomena in the linguistics of humour. Most analyses of puns to date have been taxonomic or phonological, a survey of which can be found in Hempelmann and Miller (2017). These studies describe the permissible and preferential sound transformations between a pun and its target (in terms of the types of articulatory features, the number of segments affected, their positions in the lexical and syllabic structure, etc.). Though native speakers have implicit knowledge of these transformational rules (Aarons, 2017), they must be learned or explicitly modelled in computational applications. Such models are briefly discussed in the following subsection.

Whether and how phonological features contribute to the humorousness of a pun is an open

---

[1]Not to be confused with the *target* (TA) of the GTVH.

question. Lagerquist (1980) and Fleischhacker (2005) have posited a correlation between, on the one hand, the degree of phonetic similarity between the pun and its target, and on the other hand, the "successfulness" or funniness of the pun. Hempelmann (2003a) rejects this hypothesis, basing his arguments on semantic theories of humour and on informal evidence from non-humorous pseudo-punning wordplay. More recent empirical evidence from certain forms of humorous but non-punning wordplay, however, establishes that perceived humour is a quantifiable function of entropy distance to the source word (Westbury et al., 2016). Further empirical study would be necessary to determine whether this finding also applies to puns.

Other scholarship has targeted the semantics of puns. Guiraud (1976, pp. 111–113) discusses loss of meaning as a feature of humour and observes the "defunctionalization of language" in puns. Using the GTVH as a framework, more recent studies (Attardo et al., 2002; Hempelmann, 2004; Hempelmann and Attardo, 2011) have identified the LM of puns as *cratylistic syllogism* (Attardo, 1994, Ch. 4), the notion that if meaning motivates sound, then the meaning of similar-sounding words must be similar. This line of reasoning is an example of the faulty "local" logic underpinning much humour (Ziv, 1984). Cratylism is at odds with the canonical assumption of conventional linguistics that the relation between the signifier (sound sequence) and the signified (mental concept) is arbitrary and language-specific (de Saussure, 1995, pp. 97–103). The findings of these semantic studies support my contention that humour is not suitable for processing with methods from conventional computational semantics, but rather must be treated as a special case.

2.2 Computational Linguistics and Humour

An advantage of the GTVH is that its most central notions, scripts and logical mechanisms, are amenable to mathematical and computational modelling. Attardo et al. (2002) and Hempelmann (2010) present set- and graph-theoretic models of script overlap and opposition, as well as graph-theoretic models of certain logical mechanisms. In brief, if scripts are conceptualized as sets of slot–filler pairs, then two scripts are overlapping but opposed when they have a non-null intersection, and when the complementary sets of the intersection contain subsets that are (locally) antonymic.

Though the aforementioned authors do not provide an implementation of their model, such an implementation could be realized with a knowledge base of scripts, a knowledge base of word meanings, algorithms for tagging text with reference to these knowledge bases, and an inference engine for identifying the overlapping and opposing parts.

Many of these resources, and their attendant software tools, are by now available, and of sufficient maturity, to lay the groundwork for an automated, GTVH-based interpreter of humorous text. On the knowledge base side, these resources include WordNet (Fellbaum, 1998), a lexical-semantic network storing lexicalizations and linguistic relations for over 200 000 English word senses, and FrameNet (Ruppenhofer et al., 2016), a database of some 1200 scripts covering over 13 000 English word senses.[2] On the algorithm side, state-of-the-art techniques for word sense disambiguation (Navigli, 2009) and semantic role labelling (Palmer et al., 2010) can apply WordNet senses and FrameNet scripts to raw text with 70–80% accuracy (Täckström et al., 2015; Miller, 2016). The crucial missing component here is the inference engine, which would need to identify the overlap between the tagged FrameNet scripts, and then use WordNet to find contrasting relations between the non-overlapping parts.

However, a general-purpose GTVH-based interpreter would require explicit modelling of all possible LMs, an ambitious undertaking that is far beyond the current state of the art. Restricting the interpreter to the LM of cratylism—the purview of the present paper—would be much more feasible, in particular because it could draw from the growing body of work on the computational processing of puns. This work includes attempts to computationally model the phonological properties of puns, as well as semantics-focused work on the detection and interpretation of puns. The former camp aims at producing tables of edit probabilities for sound pairs in the pun and its target, using Optimality Theory (Hempelmann, 2003a,b) or patterns learned by weighted finite-state transducers (Jaech et al., 2016). The latter camp includes studies such as Kao et al. (2016) and Simpson et al. (2019), which em-

---

[2]WordNet, FrameNet, and their counterparts for other languages have also been aligned at the word sense level (Matuschek, 2014; Hartmann and Gurevych, 2013) and combined in linked lexical-semantic resources such as EuroWordNet (Vossen, 1998), UBY (Gurevych et al., 2012), and BabelNet (Navigli and Ponzetto, 2013).

ploy Gaussian processes or information-theoretic measures to predict human judgments of the humorousness of puns, and various approaches (surveyed in Miller et al. (2017)) for sense-tagging their double meanings. While few of these systems are informed by linguistic theories of humour, analysis points to the superiority of knowledge- and rule-based approaches over rote supervised techniques.

2.3 Translation of Wordplay

The translation of humour, and more specifically of puns and other forms of wordplay, is among the most intensively studied problems in the field of translation studies (Delabastita and Henry, 1996; Delabastita, 1994; Henry, 2003; Vandaele, 2011; Regattin, 2015). Modern treatments recognize several high-level strategies for translating puns, many of which are informed by Nida's (1964) notion of dynamic translational equivalence.[3] For example, Delabastita (1996) covers the following strategies: (S_1) replace the source-language pun with a target-language pun (which may deviate from the original semantics); (S_2) substitute non-punning language that preserves one or both of the original meanings; (S_3) replace the pun with some non-punning wordplay or rhetorical device (irony, alliteration, vagueness, etc.); (S_4) omit the language containing the pun; (S_5) leave the pun in the source language; (S_6) as a compensatory measure, introduce a new pun at a discourse position where the original had none; (S_7) as a compensatory measure, introduce entirely new material containing a pun; and (S_8) editorialize: insert a footnote, endnote, etc. to explain the pun.

The choice of strategy for a given case depends on the medium of translation, the aims of the translator, the possibilities for appropriate wordplay in the target language given the source context, and the translator's ability to find and exploit these possibilities. For example, in the foreign dubbing of motion pictures for mass-market audiences, S_7 and S_8 are generally not available. Similarly, the use of S_1 can be constrained by the presence of one or both meanings in the audio or visual channel (Bucaria, 2017). The *hoarse/horse* pun of Fig. 1 exemplifies these challenges: the animal and its cough are very conspicuously shown and heard, and so must

___
[3]In contrast to formal ("word-for-word") equivalence, dynamic equivalence privileges target language solutions that aim at preserving the intention, rather than the literal meaning, of the source text. In the case of puns and other jokes, this intention is to amuse the reader in the context of the discourse.

be acknowledged in the dialogue. In the German-language dub, the translator was able to implement S_1 by having the driver say that the horse's cold had made its voice *rostig* ("rusty", but also evoking *Ross*, meaning "horse"). The translation thus preserves the original's wordplay and semantics. By contrast, the Russian dub abandons any attempt at punning, presumably because Russian lacks (or the translator could not identify) any similar-sounding synonyms for "horse" (лошадь) and "hoarse" (охрипший), nor any other pair of words for things that would make sense for the characters to be discussing in the scene. Instead, S_2 is applied, with the driver giving the non-humorous explanation, "Вчера простудился. Лошадиный кашель." ("He caught a cold yesterday. A horse cough.")

S_1 is a particular focus of the present paper. Though the literature abounds with case studies, there does not yet exist a generalized, formalized methodology for producing or explaining such translations (Delabastita, 1997). Attardo (2002) provides an important step in this direction by viewing the translation of humour through the lens of the GTVH. He argues that since the degree of perceived difference between jokes increases linearly with the salience of the KRs in which they differ (Ruch et al., 1993), translation of humour should strive to respect all six KRs. Moreover, where deviation is necessary, it should occur at the lowest level necessary for the translator's pragmatic purposes. Attardo is quick to point out that puns constitute an exception to this rule: whereas referential jokes rank LA lowest and therefore afford its translations a great deal of latitude, in wordplay, the essential features of LA are preselected by the cratylistic LM. "Translatable" puns, therefore, are those that exhibit a set of LA features in the source language that is consistent with a set of LA features in the target language, such that the pragmatic goals of the translation are met. The notion of what constitutes "consistency" of features remains an open question.

2.4 Machine Translation and Computer-assisted Translation

Though MT has made impressive strides in the last few decades, it is not yet capable of producing publication-quality output for most conventional text domains, let alone for the stylistically and semantically aberrant constructs of creative language. To date, only a scattered handful of studies have treated the topic of MT and humour. Stede

and Tidhar (1999), recognizing the unsuitability of the prevailing statistical MT paradigm for the translation of humour, propose a transfer-based architecture (Nirenburg et al., 1992) where ambiguity and its mechanisms are explicitly modelled. The architecture produces a syntactic chart of the source text and employs a measure of script opposition to identify partial analyses suitable for humorous translation. These parts are then transferred to syntactic charts in the target language, from which generation commences. However, no system implementing this architecture was ever realized, and it seems that any attempt to do so would be blocked by the infeasibility of modelling the common-sense world knowledge necessary to identify script oppositions. Farwell and Helmreich (2006) propose a separate knowledge-based translation framework that differentiates between the author's locutionary, illocutionary, and perlocutionary intents,[4] and present a case study of its application to puns. However, as with Stede and Tidhar (1999), the framework is not actually implemented as it presupposes an extensive store of real-world knowledge and beliefs that is too expensive to model computationally.

Other past work has applied MT to other forms of creative language, such as poetry. However, these studies (Greene et al., 2010; Genzel et al., 2010) focus on the preservation of surface-level constraints such as rhyme and metre, rather than semantic ambiguity as in humour. A spiritually similar constraint satisfaction approach, which targets both creative language and semantic ambiguity (but not humour), is the hidden acrostic generator of Stein et al. (2014). It relies on automatic paraphrasing, a technique intimately connected to MT (Callison-Burch, 2007). While the work is not directly applicable to the translation support task of this paper, inspiration can be taken from their optimization approach.

While much of the research in the MT community has focused on end-to-end automation of the translation process, it has been convincingly argued (Kay, 1980) that the proper role of computers is automating that which is "mechanical and routine", leaving the "essentially human" aspects to the human translator. Accordingly, there has been increasing interest in integrating information technology into traditional, manual translation workflows. Translation is by now a highly

technologized profession: electronic dictionaries, translation memories, terminology extraction systems, and concordancers are just some of the many computer-assisted translation (CAT) tools that professional translators have come to rely on (Kenny, 2011). Despite the extensive treatment creative language has received in translation studies, a survey of the available literature reveals no CAT tools that specifically support its translation. However, there exist a number of interactive, component-based CAT workbenches (Federico et al., 2014; Alabau et al., 2014; Albanesi et al., 2015) into which such support could conceivably be integrated. Some of these workbenches were designed for historical texts that pose special structural, stylistic, linguistic, and hermeneutical challenges. Their support for nonstandard text could provide useful, or at least inspirational, for handling contemporary wordplay.

3 Research Challenges

We have seen from the previous section that there exists a considerable body of foundational work on humour and wordplay in the fields of linguistics, computational linguistics, and translation studies. Linguistics provides us with semantic theories of humour that define the conditions for a text to be humorous, and with phonological models that characterize the patterns of sound changes in punning. Computational linguistics provides us with tools to automatically annotate texts with word meanings and semantic roles, to analyze lexical-semantic relationships, and to measure the semantic similarity between words and texts; the past couple of years has seen rudimentary attempts to apply these tools to the computational processing of puns. Translation studies has established a number of high-level strategies for dealing with puns, plus a wealth of case studies on the fine-grained application of these strategies. By and large, however, the past work in each of these three fields has been not been informed by work in the other two.

I contend that the time is now right to connect these separate channels of research—to start developing linguistically informed, computerized translation methodologies for dealing with the vagaries of creative language. Indeed, the need for such interdisciplinarity in the translation of wordplay has long been recognized: Delabastita (1997) acknowledged both the necessity and insufficiency of linguistic theory, arguing that pun translation mechanisms could be understood only with additional help from

---

[4]That is, how something is said, what is being said, and why something is being said (Austin, 1975).

"finer instruments... borrowed from neighbouring disciplines". And Attardo (2002) presciently remarked that "what a theory of translation really needs is a metric of similarity between meanings/ pragmatic forces". Digital versions of these instruments and metrics have become realizable only recently, with the advent of large-scale multilingual lexical-semantic resources and distributional computational semantics.

But as we have seen, past work on MT of humour and wordplay is sketchy, the only two papers on the topic (Stede and Tidhar, 1999; Farwell and Helmreich, 2006) putting forth high-level plans but no implementation or evaluation. Both papers agree that automated translation must be knowledge-based, but they do not specify how to acquire and model all the non-linguistic knowledge required to understand arbitrary instances of humour. Nor do I think that this is even possible at the present time—despite ongoing work on addressing this bottleneck (e.g., Li et al., 2016; Gordon, 2014), translation of humour requires a familiarity with discourse, cultural, and commonsense knowledge that is far out of reach of today's artificial intelligence.

Nonetheless, I believe that contemporary AI could still be made to play an important role in the translation of humorous wordplay. Rather than trying to model the entire end-to-end translation task computationally, as in MT, I argue that it is instead necessary to study how human translators approach the problem, and then provide them with tools that support rather than replace these approaches. With this "machine-in-the-loop" paradigm, language technology would be applied to only those subtasks it can perform best, such as using lexical information retrieval (i.e., searching a large vocabulary space for words matching a given set of semantic and phonological features) to generate and rank lists of pun translation candidates. Subtasks that depend heavily on real-world background knowledge and pragmatic inference—such as making the final selection from such a candidate list—would be left to the human translator.

To fulfill this vision it will be necessary to develop innovative, interactive techniques for detecting and interpreting puns in their source-language contexts, assessing the applicability of different translation strategies to a given pun, determining the amount of semantic leeway afforded to the translator, generating a set of translation candidates adhering to this semantic leeway and to phonological constraints on

punning, and dynamically exploring this candidate space. Along these lines, I envisage three major research directions:

1. Systematic and wide-ranging studies of how human translators process puns. Such studies would need to go deeper than the coarse-grained taxonomies of Delabastita (1996) and others, aiming at fine-grained models of the implementation of various translation strategies. Low (2011) provides a step in this direction by describing one method by which puns could be systematically S_1-translated, though it is not clear whether this method, or one akin to it, had been or is in common use, nor what other methods may be in use.

2. Empirical validation of the competing hypotheses (Lagerquist, 1980; Fleischhacker, 2005; Hempelmann, 2003a) concerning the relationship between a pun's humorousness and the phonetic distance between the pun and its target. Resolving this issue is a prerequisite for computationally ranking pun translation candidates by fitness.

3. Building on the above, the development of interactive, NLP-based methodologies for helping a human translator to assess whether a given pun is S_1-translatable and, if so, to implement that translation. This will necessarily involve the development and synthesis of methods based on sense annotation and semantic role labelling of puns, cross-lingual similarity of word senses and sentiments, and exploration of lexical-semantic spaces.

Though the precise form and functionality of a "Punster's Amanuensis" tool implementing these methodologies will depend on the findings of the first two research directions, the following can serve as a rough sketch. The tool would first scan the source text and flag possible instances of humorous lexical ambiguity for special attention by the translator, who can interactively confirm or reject these flags, or flag additional instances missed by the system. For each confirmed pun in the source text, the system would construct an interpretation via word sense and semantic role annotation of the pun and its context, and (in the case of multimodal source data) automatic keyword captioning (Gong et al., 2014; Vinyals et al., 2017) of any associated sound or images; these annotations would be subject

to interactive post-correction. This interpretation could then be used to identify a set of translation candidates that attempt to preserve the semantics of the original and adhere to phonological constraints on punning. These candidates could be found, for example, by looking up translations of the pun's two meanings and then searching for closely related senses in the target language whose lexicalizations have similar pronunciations. The candidates would be ranked according to various fitness measures (including phonological ones) and presented to the user in a manner that facilitates interpretation and exploration.

To expand upon this, consider the pun introduced in Fig. 1 and how the Punster's Amanuensis might help translate it into German. The tool, having detected the presence of the pun in the line, automatically interprets its two meanings and tags them with reference to their entries ("deep and harsh-sounding", with English lexicalizations *hoarse, gruff, husky*; and "solid-hoofed quadruped", with the English lexicalization *horse*) in a bilingual semantic network such as EuroWordNet. The system then looks up in the network the corresponding lexicalizations in German: *heiser* and *Pferd, Gaul, Ross*, respectively. Each possible pun–target pairing (*heiser–Pferd, heiser–Gaul, heiser–Ross*) is scored according to the phonetic similarity between the pun and the target (as determined by some phonological model of punning). For these naïve "direct" translations, the similarities are very low, indicating that they do not form valid puns in German. However, the tool can start searching the semantic neighbourhoods of the two senses for closely related senses whose German-language lexicalizations *are* similar-sounding. It might thereby arrive at *rostig* ("having a voice impaired in skill or tone by neglect") a hyponym of *heiser* that happens to have a relatively high phonetic similarity to the previously discovered *Ross*. The pun–target pair *rostig–Ross* would therefore be among the translation candidates highly scored by the system. Further candidates of this sort (*Bronchitis–Bronco*, etc.) would also be discovered and scored, either by having the system automatically expand the search space in the semantic network, or by allowing the user to manually explore it and possibly modify the search criteria.

4 Conclusion

In this paper I have surveyed the research on wordplay translation. I have observed a need for support-

ing translation technologies, but found that existing MT approaches are inappropriate and manual ones are as yet too vaguely defined to implement computationally. However, I have identified divers theories, methodologies, and resources that could be extended and integrated to produce such translation technologies. I have sketched a CAT tool encapsulating these ideas, whose exact form and functions could be refined following further studies of manual translation workflows. These studies are currently being planned, and their results (including a complete description of the resulting CAT tool) will be the subject of a follow-up paper.

References

Debra Aarons. 2017. Puns and tacit linguistic knowledge. In Salvatore Attardo, editor, *The Routledge Handbook of Language and Humor*, pages 80–94. Routledge.

Vicent Alabau, Christian Buck, Michael Carl, Francisco Casacuberta, Mercedes García-Martínez, Ulrich Germann, Jesús González-Rubio, Robin Hill, Philipp Koehn, Luis Leiva, Bartolomé Mesa-Lao, Daniel Ortiz-Martínez, Herve Saint-Amand, Germán Sanchis Trilles, and Chara Tsoukala. 2014. CASMACAT: A computer-assisted translation workbench. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (System Demonstrations)*, pages 25–28.

Keith Alan. 2001. *Natural Language Semantics*. Blackwell, Oxford.

Davide Albanesi, Andrea Bellandi, Giulia Benotto, Gianfranco Di Segni, and Emiliano Giovannetti. 2015. When translation requires interpretation: Collaborative computer–assisted translation of ancient texts. In *Proceedings of the 9th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 84–88.

Salvatore Attardo. 1994. *Linguistic Theories of Humor*. Mouton de Gruyter, Berlin.

Salvatore Attardo. 2002. Translation and humour. *The Translator*, 8(2):173–194.

Salvatore Attardo, Christian F. Hempelmann, and Sara Di Maio. 2002. Script oppositions and logical mechanisms: Modeling incongruities and their resolutions. *Humor*, 15:3–46.

Salvatore Attardo and Victor Raskin. 1991. Script theory revis(it)ed: Joke similarity and joke representation model. *Humor*, 4:293–348.

J. L. Austin. 1975. *How to Do Things with Words*. Oxford University Press, Oxford.

Chiara Bucaria. 2017. Audiovisual translation of humor. In Salvatore Attardo, editor, *The Routledge Handbook of Language and Humor*, pages 430–443. Routledge, New York, NY.

Chris Callison-Burch. 2007. *Paraphrasing and Translation*. D.Phil. thesis, University of Edinburgh.

Dirk Delabastita. 1994. Focus on the pun: Wordplay as a special problem in translation studies. *Target*, 6(2):222–243.

Dirk Delabastita. 1996. Introduction to the special issue on wordplay and translation. *The Translator*, 2(2):1–22.

Dirk Delabastita. 1997. Introduction. In Dirk Delabastita, editor, *Traductio: Essays on Punning and Translation*, pages 1–22. St. Jerome, Manchester.

Dirk Delabastita and Jacqueline Henry. 1996. Wordplay and translation: A selective bibliography. *The Translator*, 2(2):347–353.

David Farwell and Stephen Helmreich. 2006. Pragmatics-based MT and the translation of puns. In *Proceedings of the 11th Annual Conference of the European Association for Machine Translation*, pages 187–194.

Marcello Federico, Nicola Bertoldi, Mauro Cettolo, Matteo Negri, Marco Turchi, Marco Trombetti, Alessandro Cattelan, Antonio Farina, Domenico Lupinetti, Andrea Martines, Alberto Massidda, Holger Schwenk, Loïc Barrault, Frederic Blain, Philipp Koehn, Christian Buck, and Ulrich Germann. 2014. The MateCat tool. In *Proceedings of the 25th International Conference on Computational Linguistics (System Demonstrations)*, pages 129–132.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Heidi Anne Fleischhacker. 2005. *Similarity in Phonology: Evidence from Reduplication and Loan Adaptation*. Ph.D. thesis, UCLA.

Dmitriy Genzel, Jakob Uszkoreit, and Franz Och. 2010. "Poetic" statistical machine translation: Rhyme and meter. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 158–166.

Yunchao Gong, Qifa Ke, Michael Isard, and Svetlana Lazebnik. 2014. A multi-view embedding space for modeling internet images, tags, and their semantics. *International Journal of Computer Vision*, 106(2):210–233.

Jonathan Michael Gordon. 2014. *Inferential Commonsense Knoweldge from Text*. Ph.D. thesis, University of Rochester.

Erica Greene, Tugba Bodrumlu, and Kevin Knight. 2010. Automatic analysis of rhythmic poetry with applications to generation and translation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 524–533.

Pierre Guiraud. 1976. *Les jeux de mots*. Presses Universitaires de France, Paris.

Iryna Gurevych, Judith Eckle-Kohler, Silvana Hartmann, Michael Matuschek, Christian M. Meyer, and Christian Wirth. 2012. UBY – A large-scale unified lexical-semantic resource. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 580–590.

Silvana Hartmann and Iryna Gurevych. 2013. FrameNet on the way to Babel: Creating a bilingual FrameNet using Wiktionary as interlingual connection. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1363–1373.

Christian F. Hempelmann. 2003a. *Paronomasic Puns: Target Recoverability Towards Automatic Generation*. Ph.D. thesis, Purdue University.

Christian F. Hempelmann. 2003b. YPS – The Ynperfect Pun Selector for computational humor. In *Proceedings of the CHI 2003 Workshop on Humor Modeling in the Interface*.

Christian F. Hempelmann. 2004. Script opposition and logical mechanism in punning. *Humor*, 17(4):381–392.

Christian F. Hempelmann. 2010. *Incongruity and Resolution of Medieval Humorous Narratives: Linguistic Humor Theory and the Medieval Bawdry of Rabelais, Boccaccio, and Chaucer*. VDM Verlag Dr. Müller.

Christian F. Hempelmann and Salvatore Attardo. 2011. Resolutions and their incongruities: Further thoughts on logical mechanisms. *Humor*, 24:125–149.

Christian F. Hempelmann and Tristan Miller. 2017. Puns: Taxonomy and phonology. In Salvatore Attardo, editor, *The Routledge Handbook of Language and Humor*, pages 95–108. Routledge, New York, NY.

Jacqueline Henry. 2003. *La traduction des jeux de mots*. Presses Sorbonne Nouvelle, Paris.

Aaron Jaech, Rik Koncel-Kedziorski, and Mari Ostendorf. 2016. Phonological pun-derstanding. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics*, pages 654–663.

Justine T. Kao, Roger Levy, and Noah D. Goodman. 2016. A computational model of linguistic humor in puns. *Cognitive Science*, 40(5):1270–1285.

Martin Kay. 1980. The proper place of men and machines in language translation. Working Paper CSL-80-11, Xerox PARC, Palo Alto, CA, USA.

Dorothy Kenny. 2011. Electronic tools and resources for translators. In Kirsten Malmkjær and Kevin Windle, editors, *The Oxford Handbook of Translation Studies*. Oxford University Press, Oxford.

Linnea M. Lagerquist. 1980. Linguistic evidence from paronomasia. In *Papers from the Sixteenth Regional Meeting Chicago Linguistic Society*, pages 185–191.

Xiang Li, Aynaz Taheri, Lifu Tu, and Kevin Gimpel. 2016. Commonsense knowledge base completion. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1445–1455.

Peter Alan Low. 2011. Translating jokes and puns. *Perspectives: Studies in Translation Theory and Practice*, 19(1):59–70.

Michael Matuschek. 2014. *Word Sense Alignment of Lexical Resources*. Dr.-Ing. thesis, Technische Universität Darmstadt.

Tristan Miller. 2016. *Adjusting Sense Representations for Word Sense Disambiguation and Automatic Pun Interpretation*. Dr.-Ing. thesis, Technische Universität Darmstadt.

Tristan Miller, Christian F. 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*, pages 58–68.

Roberto Navigli. 2009. Word sense disambiguation: A survey. *ACM Computing Surveys*, 41(2):10:1–10:69.

Roberto Navigli and Simone Paolo Ponzetto. 2013. An overview of BabelNet and its API for multilingual language processing. In Iryna Gurevych and Jungi Kim, editors, *The People's Web Meets NLP: Collaboratively Constructed Language Resources*, pages 177–197. Springer, Berlin/Heidelberg.

Eugene A. Nida. 1964. *Toward a Science of Translating*. E. J. Brill, Leiden, the Netherlands.

Sergei Nirenburg, Jaime Carbonell, Masaru Tomita, and Kenneth Goodman. 1992. *Machine Translation: A Knowledge-based Approach*. Morgan Kaufmann, San Mateo, CA.

Martha Palmer, Daniel Gildea, and Nianwen Xue. 2010. *Semantic Role Labeling*. Morgan & Claypool, Williston, VT.

Victor Raskin. 1985. *Semantic Mechanisms of Humor*. Springer Netherlands.

Fabio Regattin. 2015. Traduire les jeux de mots: une approche intégrée. *Atelier de traduction*, 23.

Willibald Ruch, Salvatore Attardo, and Victor Raskin. 1993. Towards an empirical verification of the General Theory of Verbal Humor. *Humor*, 6(2):123–136.

Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. 2016. *FrameNet II: Extended Theory and Practice*. ICSI, Berkeley, CA.

Ferdinand de Saussure. 1995. *Cours de linguistique générale*. Grande Bibliothèque Payot, Paris.

Edwin Simpson, Erik-Lân Do Dinh, Tristan Miller, and Iryna Gurevych. 2019. Predicting humorousness and metaphor novelty with Gaussian process preference learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*.

Manfrede Stede and Dan Tidhar. 1999. Towards a humour switch for machine translation. In *Proceedings of the AISB '99 Symposium on Creative Language: Humour and Stories*.

Benno Stein, Matthias Hagen, and Christof Bräutigam. 2014. Generating acrostics via paraphrasing and heuristic search. In *Proceedings of the 25th International Conference on Computational Linguistics*, pages 2018–2029.

Oscar Täckström, Kuzman Ganchev, and Dipanjan Das. 2015. Efficient inference and structured learning for semantic role labeling. *Transactions of the Association for Computational Linguistics*, 3:29–41.

Jeroen Vandaele. 2011. Wordplay in translation. In Yves Gambier and Luc van Doorslaer, editors, *Handbook of Translation Studies*, volume 2, pages 180–183. John Benjamins.

Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. 2017. Show and tell: Lessons learned from the 2015 MSCOCO image captioning challenge. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):652–663.

Piek Vossen, editor. 1998. *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Springer.

Chris Westbury, Cyrus Shaoul, Gail Moroschan, and Michael Ramscar. 2016. Telling the world's least funny jokes: On the quantification of humor as entropy. *Journal of Memory and Language*, 86:141–156.

Avner Ziv. 1984. *Personality and Sense of Humor*. Springer.