Computational Analysis and Generation of Slogans

Khalid Alnajjar

Helsinki November 5, 2018

UNIVERSITY OF HELSINKI
Department of Computer Science

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Studieprogram — Study Programme
Matematisk-naturvetenskapliga fakulteten	Datavetenskap

Tekijä — Författare — Author

Khalid Alnajjar

Työn nimi — Arbetets titel — Title

Datorbaserad analys och generering av slagord

Ohjaajat — Handledare — Supervisors

Prof. Hannu Toivonen

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Magisteruppsats	11 november 2018	77 sidor + 8 bilagor

Tiivistelmä — Referat — Abstract

I reklam används sloganer för att förbättra återkallandet av den annonserade produkten av konsumenter och skilja den från andra på marknaden. Att skapa effektiva slagord är en resurskrävande uppgift för människor. I denna avhandling beskriver vi en ny metod för att automatiskt generera sloganer, med tanke på ett målkoncept (t ex bil) och en adjektivsegenskap för att uttrycka (t ex elegant) som input. Dessutom föreslår vi en metod för att generera nominella metaforer med hjälp av en metafor-tolkningsmodell för att möjliggöra generering av metaforiska slagord. Metoden för att generera sloganer extraherar skelett från befintliga sloganer, så fyller det ett skelett med lämpliga ord genom att använda flera språkliga resurser (som ett förvar av grammatiska och semantiska relationer och språkmodeller) och genetiska algoritmer samtidigt som man optimerar flera mål såsom semantiska relateradhet, språkkorrigering och användning av retoriska enheter.

Vi utvärderar metaforen och slogangenereringsmetoderna med hjälp av en tänktalkoplattform. På en 5-punkts Likert-skala ber vi online-domare att bedöma de genererade metaforerna tillsammans med tre andra metaforer som genererades med andra metoder och visa hur bra de kommunicerar den eftersökta betydelsen. Slogangenereringsmetoden utvärderas genom att be crowdsourced-domare att bedöma genererade sloganer från fem perspektiv, vilka är 1) hur bra är sloganet relaterat till ämnet, 2) hur korrekt är sloganets språk, 3) hur metaforiskt är sloganet, 4) hur engagerande, attraktivt och minnesvärt är det och 5) hur bra är sloganet överlag. Dessa frågor är utvalda för att undersöka effekterna av relateradhet till produkten och den markerade egenskapen, användningen av retoriska anordningar och språkets korrekthet på den övergripande uppskattningen av slogan. På samma sätt utvärderar vi befintliga sloganer som har skapats av äkta människor. Baserat på utvärderingarna analyserar vi metoden som helhet tillsammans med de enskilda optimeringsfunktionerna och ger insikter om befintliga sloganer. Resultaten från våra utvärderingar visar att vår metaforgeneringsmetod kan producera lämpliga metaforer. För slogangeneraren bevisar resultaten att metoden har varit framgångsrik i att producera minst en effektiv slogan för varje utvärderad input. Ändå finns det utrymme för att förbättra metoden, som diskuteras i slutet av avhandlingen.

ACM Computing Classification System (CCS):
Computing methodologies → Artificial intelligence → Natural language generation
Computing methodologies → Artificial intelligence → Language resources

Avainsanat — Nyckelord — Keywords

Slogan Generation, Computational Creativity, Natural Language Processing and Generation

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — övriga uppgifter — Additional information

Avhandling för under programmet Algoritmer, Data Analytics och Maskininlärning

Contents

1 Introduction

In advertising, it is essential to construct expressions wisely in order to convey a precise message as persuasive and clear as possible, and also because of the limiting constraints when advertising (e.g. character limit, budget, and time).

Slogans are memorable short phrases that express an idea or product, and are frequently used in advertising and branding. Slogans are commonly used in advertising campaigns to enhance the recall of the product by customers and distinguish it from others.

When encountering the phrase "Just Do It.", a concept immediately pops up in our minds, *Nike*. It is fascinating how a phrase is firmly associated with a brand in our minds despite how it is perceived by the audience (e.g. as an imperative clause telling the receiver to do sports or buy *Nike*'s products). This effect highlights the significance of slogans in advertising.

The objective of this thesis is to propose a method for automatically generating a list of slogan candidates, given a target concept (e.g. car) and an adjectival property to express (e.g. elegant) as input. The intended use-case of the method is assisting advertising professionals during brainstorming sessions by dynamically suggesting slogans tailored to their needs. Suggested slogans by the method are not expected to be used as they are generated in production, but rather some human refinement is needed.

To fulfil the objective of this thesis and build a model that produces advertisable slogans, we commence by reviewing the current literature regarding analysing and generating slogans and other creative expressions. Next, we manually and computationally analyse existing slogans crafted by professionals of various brands. We then implement a model for generating slogans based on our analysis and the reviewed literature on slogan generation.

In the literature review, we summarise existing research on analysing and generating slogans, such as the work described by Mcquarrie and Mick [36] and Özbal, Pighin, and Strapparava [39], respectively. The research concerning slogans is scarce, particularly computational analysis. Thus, our literature review of slogan analysis comprises literature of both manual and computational analysis of slogans. While reviewing computational methods for generating slogans, we highlight how they differ in achieving the task on aspects such as linguistic processing (e.g. ensuring correct grammar and measuring relatedness) and the algorithmic process of generating slogans (e.g. selecting or producing slogan candidates and ranking them).

Following our literature review, we annotate a random sample of man-made slogans. The goal of this analysis is to discover linguistic characteristics that are used regularly in them (e.g. usages of rhetorical devices). We also perform a semi-manual analysis where we analyse the same slogans using computational tools and then audit the results manually.

In addition to the manual analysis, we define computational measurements to computationally bulk-analyse existing slogans. Examples of such measurements are employing the *CMU pronouncing dictionary* to measure consonance and assonance, and using word frequencies to detect how common words in slogans are. Moreover, we utilise Natural Language Processing (NLP) resources to parse and analyse the natural language of slogans, e.g. tokenize them and obtain their syntactical structures. We calculate semantic cohesion and relatedness to the input by using semantic associations measures.

A computational approach is then proposed by us for generating slogans. The approach is based on the reviewed computational approaches in addition to any computational methods we find applicable for the task and reflect our analysis. A key component of our proposed method is a method for generating metaphors, which are used in our slogan generation method to generate metaphorical slogans.

We evaluate the proposed method, including the metaphor generator, by running crowd-sourced surveys. Additionally, we run a similar survey on man-made slogans to present further analysis and results, with respect to computer-made slogans. We then discuss the results and provide suggestions for future work.

Our contributions in this thesis focus on studying slogans from a linguistics perspective and how we can produce slogans computationally, especially metaphorical ones. Moreover, it contributes by scrutinising the method, as well as existing slogans, from multiple angles revealing fresh insights.

Part of the work presented in this thesis has resulted in a conference publication by Alnajjar, Hadaytullah, and Toivonen [3], which describes our metaphor and slogan generation, and some of the evaluation results. The method proposed in this thesis has been adapted by Alnajjar and Hämäläinen [1] to generate cultural satire movie titles, a different type of creative expressions. In addition to the conference publications, we have produced and publicly released a dataset containing the 12 million most frequent English grammatical relations and their frequencies [4]. These publications emphasise the contributions of this thesis.

This thesis is structured as follows; we start by describing the issue which this thesis is addressing in section 2, where we also provide preliminaries regarding the topic, define the research questions and the problem of generating slogans computationally formally from a computational perspective, and motivate this research.

Then, we review related work on slogan analysis and generation in section 3. In the analysis, we cover automated and manual studies of linguistic characteristics of advertising slogans, whereas the literature of slogan generation is focused on the computational aspect only. As our slogan generator uses the output of our method for generating metaphors, we briefly summarise some of the primarily related work on metaphor generation.

Section 4 details the corpora of slogans used in our work. In section 5, “Manual Analysis of Slogans”, we manually analyse linguistic characteristics of randomly selected existing slogans of brands to discover and highlight important linguistic characteristics, and to categorise them based on their usage of rhetorical devices. Following that, section 6 reveals further linguistic characteristics by employing automatic text analysis tools and, then, manually verifying the results.

Computational analysis and generation of slogans fall under section 7. In it, we state the computational and linguistic resources used in the method and define various, novel and existing, functions (e.g. rhyming, number of characters and semantic relatedness), which are examined against the entire corpora of slogans. Following the analysis, we provide our findings. Inspired by the findings, we elucidate our computational generation methodology.

We conduct experiments to evaluate our method in section 8. The experiments examine the method for generating metaphors and slogans, whether they are computer-generated or man-made. The results of the experiments are given in the same section. Thereupon, we discuss the method and our findings in section 9 followed by section 10 where we conclude the thesis and promote future work.

2 Problem

Coming up with successful slogans is challenging, for humans and machines. The goal of this thesis is to produce a model that generates advertising slogans for concepts (e.g. brands, ideas and products) computationally.

In this section, we briefly explain the advertising and linguistic prerequisite preliminaries that are related to slogans. Furthermore, we motivate the work conducted in this thesis. Following that, we define the problem of producing slogans from a computational perspective and the research questions we focus on answering in this thesis.

2.1 Preliminaries

Slogans are a form of advertisements. Advertisements are ubiquitous, we see them on daily basis all the time everywhere. Richards and Curran [41] define advertising as "a paid, mediated, form of communication from an identifiable source, designed to persuade the receiver to take some action, now or in the future".

Slogans, taglines, mottoes are very similar to the extent that they are considered synonyms. Slogans and taglines are often used interchangeably; however, slogans are made for an advertising campaign whereas taglines are employed as an identifiable phrase for the brand. In other words, a slogan is made for one or more advertising campaigns but a tagline is typically made once for the lifetime of the company. On the other hand, mottoes are sayings that represent a group's (e.g. corporate, political, and religious) vision such as *Google*'s "Don't be evil"[1]. In this thesis, we consider all retrieved expressions online as slogans, given the similarities they have and the difficulties in distinguishing them.

Slogans are a complex concept. They are utilised not only by companies (e.g. *Nokia*'s "Connecting People"[1]) but also by political parties, Barack Obama's 2008 campaign slogan "Yes We Can"[2]; religious groups, *United Methodist Church*'s "Open hearts, Open minds, Open doors"[1]; universities, the motto of *Aalto University School of Arts, Design and Architecture* "Pro Arte Utili" (in English: "For useful abilities")[3]; and many others. Despite of the field, slogans generally are used in advertising campaigns.

We define a slogan, from an advertising perspective, as a concise, advertisable, and autonomous phrase that expresses a concept (e.g. an idea, product, or entity) which will be frequently repeated and associated with it. Elements of advertisability

[1] The example is taken from https://en.wikipedia.org/wiki/List_of_mottos

[2] Obtained from the Wikipedia article of the list of U.S. presidential campain slogans: https://en.wikipedia.org/wiki/List_of_U.S._presidential_campaign_slogans

[3] Based on: https://en.wikipedia.org/wiki/Aalto_University_School_of_Arts,_Design_and_Architecture

include creativity, catchiness (i.e. draws attention), memorability (i.e. easy to memorize and recall), clearness (i.e. does not cause confusion), informativeness (i.e. has a message), and distinctiveness (i.e. unique) [10]. Creating components of advertisements, such as slogans, while maintaining these elements manifests the difficulty of the task, for humans and computers.

Creativity, an element of advertisability, is defined by Amabile [5] as "the production of a novel and appropriate response, product, or solution to an open-ended task". The task of producing slogans is an open-ended task as it is neither trivial nor has a single solution. Regarding the solution produced for the task, a slogan in our case, it must be practical, solving the task and new. Furthermore, it must be appreciated and seen as creative by judges. All the aspects of Amabile's definition of creativity are found in slogans (from the task itself to the process of creating them and the necessity of appreciating the final slogan by the consumers and other advertisers).

Language is a device for communication, and slogans are a type of such device as they convey a message to the receiver, usually a persuasive message about a concept. Like poems, slogans naturally have a stylistic language. Stylistic language is concerned with *how* a message is expressed rather than its content. Rhetorical devices –figures of speech– are examples of stylistic language. They exploit the listeners' knowledge of the language to persuade them by redirecting their knowledge into the speaker's intended path. Linguistic creativity can be demonstrated by utilising rhetorical devices aptly.

Many slogans employ rhetorical devices. For instance, *Yellow Page*'s slogan "Let your fingers do the walking." uses personification expressing fingers as entities capable of walking[4]. Previous research suggests that slogans employing rhetorical devices tend to be more favoured and remembered by consumers [40]. Moreover, different rhetorical devices in slogans have various effects on consumers. For instance, Burgers et al. [7] suggest that slogans containing conventional metaphors are liked and considered more creative than slogans containing irony.

Mcquarrie and Mick [36] have analysed rhetorical figures in advertising language used in printed ads. They have proposed a classification of rhetorical figures in advertising which we will utilise in our analysis. Figure 1 illustrates the hierarchical taxonomy of rhetorical devices proposed by Mcquarrie and Mick.

The hierarchy consists of three levels, i) figuration, ii) figuration mode and iii) rhetorical operation. Figuration, the first level, aims to distinguish rhetorical figures from

[4] Examples in this section are from the collected corpora in section 4, unless otherwise specified.

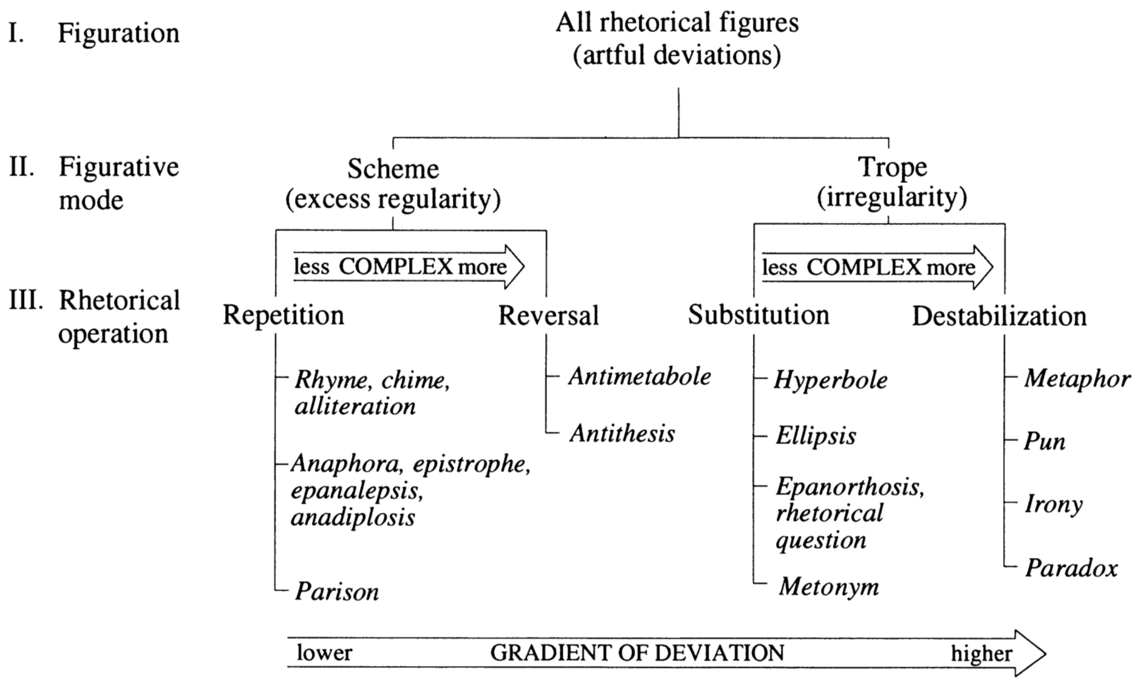


Figure 1: The taxonomy of rhetorical figures in advertising as proposed by Mcquarrie and Mick [36]

literal expressions. Separating figurative slogans from literal ones can be accomplished by considering any expression that is an artful deviation from the listener's expectation as rhetorical [9]. The deviation is based on the design of the expression, not the content. A standard that can be used to measure such deviation, listener's expectation, is whether the meaning of the words is as defined in the dictionary. For instance, the metaphor in *Johnson & Johnson*'s Band-Aids slogan, "Say hello to your child's new bodyguards", violates such expectation as bodyguards are expected to protect an important or famous person, not children [36]. On the other hand, the slogan "Giving insurance solutions 'from A to Z' " by *Allianz* simply describes the product provided by the company without using any figurative language; hence, it is a literal slogan. Expressions can have a different degree of deviation. However, the deviation cannot be a faulty diction or bad grammar as these are errors.

The second level of the hierarchy –figurative mode– is concerned of categorising rhetorical devices into schemes (e.g. rhymes such as *Chevrolet*'s "Eye it. Try it. Buy it.") and tropes (e.g. metaphors such as *Toshiba*'s slogan "In touch with tomorrow"). Tropes are figures of speech that are on the meanings level, whereas schemes are on syntax, words, letters and phonetics level.

The last level in the hierarchy is rhetorical operations. Rhetorical operations consist of four types, which are i) repetition, ii) reversal, iii) substitution and iv) destabilization, ordered from low to high gradient of deviation respectively. Repetition and reversal are types of schematic mode, whereas substitution and destabilization are types of tropic mode.

Repetition indicates that parts of the expression (e.g. words, letters, and sounds) are repeated while maintaining the meaning of the expression, for example, *Early Learning Centre*'s slogan "Special toys, special prices" depicts the repetition of words. Rhymes, chimes, and alliteration are examples of repetition is on the sound level (e.g. *Rayon*'s slogan "It's the knit with the fit where you sit." and *Pepsi*'s "More bounce to the ounce"), whereas anaphora, epistrophe, epanalepsis and anadiplosis are on word level (e.g. *IBM*'s slogan "IBM. Computers help people help people."). On the other hand, parison is a rhetorical device that repeats the structure of a phrase (e.g. *Alcombe Veterinary Surgery*'s slogan "We listen to you, we talk to the animals.").

Reversal, a schematic operation too, is concerned of reversing the syntax as in antimetabole (e.g. repeating words in reversed order such as *Bridgestone Dueler*'s slogan "First you drive it. Then it drives you.") or semantics as in antithesis (e.g. using opposite terms such as *Church's Chicken*'s slogan "Big Pieces, Little Prices") in a phrase.

Tropic figures have two types, substitution and destabilization. In both types, the receiver is expected to settle the meaning of the slogan to understand the conveyed message. However, the main difference between the two types is that substitution tends to have one possible alternation but destabilization can have multiple possible interpretations.

In substitution rhetorical operations, there are four dimensions of rhetorical figures: i) stating an extreme claim (exaggerating as in hyperbole or understating as in litotes), ii) omitting a word/phrase from a sentence as in ellipsis, ii) implying weak/strong assertions as in epanorthosis and rhetorical question, and iv) referring to a given concept by a closely associated element with it as in metonymy.

Destabilization, also, has four rhetorical dimensions (metaphors, puns, irony, and paradox). The following four paragraphs introduce these four concepts.

Metaphors are figurative expressions consisting of two concepts, a tenor and a vehicle following Richards [42] terminologies, where some properties get highlighted or attributed to the tenor from the vehicle. For example, in *Oakmont Bakery*'s slogan

"We create delicious memories" the tenor (memories) is represented by a common property of the vehicle (food) which is delicious.

Puns exploit polysemy or homonymy to draw receiver's attention by creating a semantic disambiguation. For example, the pun in *Skil*'s "Do it right. Do it with Skil." exploits the phonetic similarities between the word *skill* and the brand *Skil*.

Irony happens when the slogan means the opposite of what is said. *Acuvue*'s slogan for their disposable contacts "We spent years developing this incredibly comfortable contact lens, and this is how you treat it" (advertised with a picture of a lens flicked away) is an example of an ironic advertisement [36]. In this example, the true meaning of the phrase is to reproach the audience for throwing their contacts after all the hard-work *Acuvue* has put in producing them. Notwithstanding, the intended meaning is the opposite (i.e. we have spent all these years so you can treat these lenses in this way).

Paradox occurs when a given phrase is self-contradictory. For instance, consider the slogan of the *U.S. Army* "Some of our best men are women.". As it is impossible for men to be women at the same time, this slogan raises a paradox conflict for the receiver.

Some rhetorical devices are not covered in the taxonomy proposed by Mcquarrie and Mick (e.g. affixation, similes, personification …etc). Although, such rhetorical devices can still fall under the suggested rhetorical operations (e.g. considering similes and personifications as destabilization operations). As we are following their taxonomy, we also follow their definitions of the figurative devices while classifying slogans to lessen confusions. For instance, based on their definitions, chime is "keywords in a phrase begin with identical sounds or letters" and alliteration is "three or more repetitions of constants.". These definitions are not strictly correct, as chime should be substituted with alliteration and consonance should replace alliteration, based on the Oxford University Press (http://www.oed.com).

A slogan may contain multiple rhetorical operations to demonstrate greater creativity. For instance, *Skil*'s slogan "Do it right. Do it with Skil." contains both repetition and destabilization rhetorical types.

Studying the meaning of slogans is important as it would not only help in comprehending slogans but also aid in building a computational model that assembles expressions conveying meaning similar to man-made slogans. Depending on the context, some expressions can have multiple interpretations. Semantics and pragmatics are sub-fields in linguistics which study the meaning of expressions in the language. The two coming paragraphs describe them in short.

Semantics focus on studying what words denote and mean. We concentrate our semantics research to lexical semantics, more specifically on how lexicons in slogans are associated with the concept (product or brand) and its category.

Utterances can be ambiguous as their meaning would vary from context to other. And, sometimes, more is understood than what is stated in them. Pragmatics studies the meaning of linguistic expressions from these aspects. Slogans fit properly under these aspects as they are usually advertised without context (rather the context is assumed by the listener) and the audience interprets and understands them in a wider sense. One way of understanding their meaning as a receiver is by analysing them from a specific pragmatic theory called speech acts.

Speech acts theory is based on the idea that the meaning of utterances depends on their use in context, rather than the words used in them. The theory states that spoken utterances can fall under three types. Illocutionary, the type we are interested in, considers utterances produced by the speaker as actions that are understood based on the speaker's intentions. Moreover, it states that these utterances can alter the social reality, called performative utterances.

Searle [43], suggests that there are five classifications of illocutionary utterances, which are representatives; (e.g. stating facts and descriptions), directives; (e.g. requesting and demanding), commissives; (e.g. promising), expressives; (e.g. giving excuses and congratulating), and declaration; (e.g. firing an employee, pronouncing a president and a couple as husband and wife). We focus this thesis on semantic analysis only. However, some research has been carried out on analysing the pragmatics of slogans [55].

Advertising slogans tend to contain positive words [14] which would give the receiver a positive feeling about the brand. Consequently, it is essential to analyse the sentiment of existing slogans and employ sentiment analysis when constructing slogans.

2.2 Motivation

This section motivates the work of this thesis. We start by motivating the work from a computer science perspective. Thereafter, we motivate it from various domains such as finance, marketing, business and news.

Constructing successful slogans is a challenge for humans as it requires multiple criteria to be met, as we have described in section 2.1. Creativity is one of these criteria which is also challenging for computers; hence, the increased research on creativity in the computer science field, computational creativity.

Solving a creative task, such as producing advertisable slogans, by computational means is motivated by computational creativity. Computational creativity, based on the definition of Colton and Wiggins [8], is "the philosophy, science and engineering of computational systems which, by taking on particular responsibilities, exhibit behaviours that unbiased observers would deem to be creative". As branding and advertising professionals are generally considered creative, a computer with a similar ability to produce slogans should also be considered creative. Moreover, in building a model for generating slogan candidates tailored to the user's desire, we provide professionals with the ability to collaborate with computers to produce creative results more efficiently.

Our research adds to the fields of Natural Language Processing (NLP) and Natural Language Generation (NLG) in two clear directions. The first is a general contribution in which we computationally process and generate short natural language, slogans; in addition to suggesting further improvements based on our take on achieving the task. The second contribution is focused on figurative language. Figurative language is used intensively in slogans. We will study the use of figurative language in existing slogans to obtain a degree of understanding of the linguistic creativity of humans used in slogans. This work will address generating and evaluating some figurative devices in short expressions.

In terms of financial impact, in 2016, €345 million were spent on online advertising in Finland only[5]. Online advertising is just a drop in the ocean, and this figure symbolises how gigantic the advertising market is. Slogans are an essential part of advertising campaigns and automating their generation would reduce the costs of advertising campaigns and make creating them more efficient.

Slogans change over time and typically are not fixed for advertising campaigns [29]. Brands may change their slogans, for instance, to target a certain audience, provide a new persuasive selling statement for a given product, or reflect changes in the company's values. Mathur and Mathur [34] have found firms that change their

[5] According to Interactive Advertising Bureau Europe 2016 report https://www.iabeurope.eu/wp-content/uploads/2017/06/IABEurope-AdExBenchmark-2016-full-report.pdf

slogan seem to have positive effects on their market value. Tom and Eves [45] have found that advertisements containing rhetorical figures are more persuasive and have higher recall in comparisons to slogans that do not utilise rhetorical figures. Additionally, a research conducted by Reece, Van den Bergh, and Li [40] suggests that recalling a slogan relies largely on the slogan itself, not on the advertising budget, years in use or themes. Such continues change of slogans can benefit from a slogan generator.

A brand's identity is its name, logo, and tagline (treated as a slogan in this thesis, refer to section 2.1). Many researchers have studied the importance of using slogans. For instance, Boush [6] argues that slogans can prime customers of a potential new brand product as an extension of the current branded products of the same category. Such priming effect can vary depending on how general the slogan is; therefore, the level of generality of the slogan should be well selected to increase the intended priming effect.

Persuasive messages are not only used in slogans, but news headlines also employ it a lot to encourage the audience to read the article [18]. Gatti et al. [20] have demonstrated how well-known expressions (such as slogans) can be utilised to produce interesting news headlines. As a result, a creative slogan generator is not limited to the advertising domain but also can be employed in various domains where persuasive short linguistic expressions are used.

2.3 Definition

In this section, we divide and formulate the problem of creating human-like slogans computationally as research questions. We explain why these questions are chosen, and how we are going to answer them.

As we strive to build a generation model that reflects how advertising professionals craft slogans, our first set of research questions are focused on discovering and understanding the characteristics of human-made slogans.

The first research question in the set is RQ1: *What are the linguistic characteristics of human-made slogans?*. The linguistic characteristics of slogans that we are interested in investigating are (RQ1.1) phonology, (RQ1.2) syntax, (RQ1.3) semantics and (RQ1.4) sentiment. We will study these characteristics to highlight the linguistic features that slogans tend to have by utilising computational resources (e.g. Natural Language Processing and semantic relatedness models) on existing slogans when feasible; otherwise, they will be analysed manually.

Our second research question, RQ2.1, is *Which figurative devices are commonly used in slogans made by professionals?*. Then, we look at *which figurative devices tend to be frequently associated with other devices?*, RQ2.2. We intend to provide answers to these questions by manually analysing human-made slogans as the current state of computational methods for analysing figurative language in texts is not advanced. Answering these questions would assist in answering a bigger question, RQ2.3 *how can we computationally generate figurative devices (such as metaphors) and expressions containing them?*. Due to the scope of the thesis, we focus on the figurative device most seen in slogans.

Moving from research questions focusing on slogan analysis and figurative devices to the computational generation of slogans, we study RQ3 *how can we generate slogans and evaluate certain linguistic criteria in them automatically?*. While answering the research question RQ3, we look at RQ3.1 *how can a machine create similar syntax as in man-made slogans?* and RQ3.2 *what are the requirements for semantic knowledge to generate slogans that convey meanings similar to existing ones?*. RQ3.3 looks at *which linguistic and computational resources and tools are needed to accomplish the task of generating suitable and metaphorical slogans?*. By reviewing the literature on Natural Language Generation (NLG), more specifically on the generation of slogans and figurative language, we will understand the current state-of-the-art in achieving such tasks. We, also, will investigate and conduct trial and error experiments while building the model for generating slogans.

After finding out a way of generating slogans, we investigate RQ4 *whether the slogans generated by our model considered suitable?*. In addition to that, we evaluate the generated slogans from other aspects, such as the correctness of the language and their catchiness. We will answer this question by running crowdsourced surveys.

To share further justifications and insights on our method and its output, we raise the following research question: RQ5.1 *what compelling outcomes can we tell from our experiments?*. In addition to that, we juxtapose computer-made slogans with human-made slogans from different aspects to find out RQ5.2 *how the results obtained from our evaluations on computer-generated slogans relate to man-made slogans?*. This research question explores the performance of our method and points out future directions to improve the method, to reacher a similar level to man-made slogans.

Computational definition In this section, we define the task of generating slogans from a computational perspective. Given an input concept, T, (e.g. car) and an adjectival property P to express (e.g. elegant), the method produces a list of weighted slogan candidates, $\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3 \ldots \mathcal{E}_n\}, n \geq 1$. Each slogan, $\mathcal{E}$, is represented as a sequence of tokens (i.e. words, numbers or punctuation marks, excluding spaces) $\{t_1, t_2, t_3 \ldots t_m\}, m \geq 1$. The main task of the method is to construct slogan candidates that contain certain desired criteria of slogans (e.g. relatedness to the input, usage of rhetorical devices and correct grammar) and output the top slogan candidates to the end-user. As a result, it is necessary for the method to internally evaluate the produced candidates based on the desired criteria of slogans.

As we address the task of generating metaphors (a figurative device commonly used in slogans) in our method, we define the metaphor generation task as follows. Given a tenor/concept T and an adjectival property P, the generator produces vehicle candidates, $V = \{v_1, v_2, \ldots v_i\}$. A vehicle highlights the adjectival property P in T when perceived metaphorically. Two examples of a vehicle candidate for expressing that a *computer* is *creative* are *poet* and *music*. It is worth noting that the interpretations of these two vehicles could be different. The former might highlight that the computer is an entity producing creative output, whereas the latter might highlight that the computer is a creative work. Nevertheless, both metaphors could highlight the property creative.

3 Related Work

While the research on computational analysis and generation of creative expressions, such as slogans, is rapidly increasing, it is still scarce. Nevertheless, a few researchers have tackled these topics. In this section, we describe the related work on analysing slogans from a linguistics perspective. Thereafter, we provide a brief summary of the literature on generating nominal metaphors as we generate metaphors in our method. Lastly, we explain the current computational approaches for generating slogans and other creative expressions.

3.1 Linguistic Analysis of Slogans

Slogans are a special type of language, like poetic language, as they have their own linguistic characteristics.

Current research on slogan analysis is concentrated on manually analysing the linguistic characteristics, mainly on their usage of rhetorical figures.

Reece, Van den Bergh, and Li [40] have analysed linguistic characteristics of slogans, in addition to other characteristics such as their themes, to find out how they affect receivers in recalling the brand. Their study indicates that utilising some linguistic devices has indeed affected the recall of the brand. The top eight slogans with high-recall contained the following linguistic devices: 1) self-reference (i.e. having the brand name in the slogan), 2) alliteration, 3) parallel construction (which is anaphora in the taxonomy by Mcquarrie and Mick [36]), 4) metaphor and 5) well-known phrase. The authors have also noticed that the slogan with the highest number of correct brand-identifications made use of rhymes. As a result, these linguistic devices have a significant influence on recalling the brand. Albeit, some of the frequently found linguistic devices in slogans did not have such outstanding influence, e.g. puns.

Inspired by the analysis and taxonomy of linguistic devices used by Reece, Van den Bergh, and Li [40], Miller and Toman [37] have manually analysed slogans from various linguistic perspectives, focusing on rhetorical figures and covering other linguistic devices. Their research shows that linguistic devices existed in $\approx 92\%$ of 239 slogans, out of which 80% and 42% were schematic and tropic rhetorical devices, respectively. Additionally, the two most common rhetorical devices which were found in figurative slogans are phonetic and semantic devices, covering 87% and 37% of them. Some phonetic devices appeared more than others, e.g. both consonance and assonance devices occurred in 59% of figurative slogans whereas 32% and 4% of them had alliteration and rhyming respectively. Other linguistic devices analysed by the authors are syntactic, orthographic and morphological devices which appeared in less than thirty slogans.

A similar manual analysis was conducted by Dubovičienė and Skorupa [15]. Their results also demonstrated that slogans use rhetorical devices frequently, especially figurative language and prosody. However, the distribution of the individual rhetorical devices does not match the one by Miller and Toman [37], which could be due to the different sources of slogans used during the analysis.

Other linguistic characteristics were studied too. For instance, Șimon and Dejica-Carțiș [55] have looked at the usage of speech acts in 84 ads and found that some speech acts (e.g. information, directions and assertion) are more frequently found than others (e.g. claiming and thanking). Notwithstanding, we focus our manual analysis of slogans in this thesis on figurative devices.

In terms of computational linguistic analysis of slogans, Dowling and Kabanoff [14] have analysed 240 advertising slogans by employing content analysis. They have applied statistical analysis to discover which categories are commonly used in slogans. In their analysis, they had 22 categories of terms. For instance, the category *Economic 1* would contain terms such as *business, commerce, trade* ...etc, and the category *Time* would include terms such *past, future, year* ...etc. Their results show that the categories *strength, active, positive, and economics* are the most frequent in slogans. An interesting insight of their results is that there was no significant link between the categories of terms used in slogans and the main nature of services provided by their corresponding companies. Dowling and Kabanoff [14] state that slogans tend to have words focusing on one to two categories. The authors have also clustered these slogans into 5 groups and followed that with the same analysis to report which categories are the most representative of each group. Moreover, they argue that slogans should be descriptive and provoke positive emotional feeling with the description given.

To our knowledge, the research conducted by Dowling and Kabanoff [14] is the only automated linguistic analysis of slogans. In this thesis, we conduct additional analysis on slogans from various aspects such as prosody, semantic relatedness, sentiment analysis, length of slogans and word infrequencies.

3.2 Computational Generation of Metaphors

We will provide some literature on metaphor generation, more specifically generating nominal and conceptual metaphors for a given tenor. This is because, in this thesis, we also introduce a method for generating nominal metaphors.

The approach proposed by Xiao and Blat [52] is focused on generating metaphors for pictorial advertisements. Their approach utilises multiple knowledge bases, e.g. word associations and common-sense knowledge[6], to find concepts with high imageability. The found concepts are then evaluated against four metrics, which are affect polarity, salience, secondary attributes and similarity with the tenor. Concepts with high rank on these measures were considered apt vehicles to be used metaphorically.

Metaphor Magnet[7], a web-service built by Veale and Li [50], generates and inter-

[6] ConceptNet: http://www.conceptnet.io
[7] http://ngrams.ucd.ie/metaphor-magnet-acl/

prets metaphors by observing the overlap of stereotypical properties between concepts. The stereotypical associations are retrieved from Google 3-grams having the pattern "a/n *ADJ NOUN*", such as "a scary clown". They are, then, validated by querying the web-specific linguistic patterns (e.g. "as *ADJ* as a *NOUN*") and retaining relations appearing at least once on the Internet. To ensure a high-quality associations, the authors have manually pruned out wrong associations. The remaining associations are expanded and used to generate metaphors by measuring their aptness based on the associations shared with the tenor and vehicle.

Galvan et al. [19] generated metaphors by using a web service, *Thesaurus Rex* [51], that provides categorisations of concepts and adjectival properties associated with them. Their approach starts by retrieving top 40% categories of the input tenor. It then selects an adjectival property, at random, that is associated with the tenor. Thereafter, it sends another query to the web service to obtain categories associated with the previously selected property. A category matching the retrieved categories of the tenor is selected. Finally, it creates a metaphor by finding a concept falling in the selected category which is also strongly associated with the selected property.

In contrast to the reviewed metaphor generation methods, our method employs a metaphor interpretation model to identify apt metaphors.

3.3 Computational Generation of Slogans

Strapparava, Valitutti, and Stock [44] proposed a creative function for producing advertising messages automatically. Their approach is based on the "optimal innovation hypothesis" [21]. The hypothesis states that the optimal innovation is reached when novelty co-exists with familiarity, which encourages the recipient to compare what is known with what is new, resulting in a pleasant surprise effect. The approach proposed by the authors utilises semantic and emotional relatedness along with assonance measures to find interesting candidates of words to substitute some existing words in human-made familiar expressions.

Özbal, Pighin, and Strapparava [39] have introduced a framework, called *BrainSup*, for creative sentence generation. The framework generates sentences such as slogans by producing expressions with content semantically related to the target domain, emotion and colour, and some phonetic properties. The generated expressions must contain keywords that are input by the user. Using syntactical tree-banks of existing sentences as sentence skeletons and syntactical relations between words as constraints for possible candidate fillers, Özbal, Pighin, and Strapparava have employed beam search to greedily fill in the skeletons with candidates meeting the desired criteria.

Using *BrainSup* as a base, Tomašič, Żnidaršič, and Papa [48, 47] have proposed an approach for generating slogans using genetic algorithms instead of beam search. Moreover, their evaluation criteria were different from *BrainSup*'s evaluation. Their work demonstrated how it is possible to automatically generate slogans without any user-defined target words by extracting keywords from the textual description of the target concept.

Case-based reasoning (CBR) was also used in generating slogans. The approach proposed by Żnidaršič, Tomašic, and Papa [54] employs case-based reasoning where actual slogans written by humans (not their syntactical skeletons) were reused with some modifications in a different context as a new slogan. The approach commences by retrieving related slogans to the textual description of the input concept using semantic similarities. Slogans are then transformed by replacing content words in them with words from concept's description while satisfying existing part-of-speech (POS) tags.

Regarding figurative language generation, *Figure8*, a system proposed by Harmon [25], generates metaphorical sentences. Five criteria were considered in the generation process, namely: clarity, novelty, aptness, unpredictability, and prosody. The system selects a tenor and searches for a suitable vehicle to express it. Thereafter, it composes sentences to express the metaphor by filling templates of metaphorical and simile expressions.

An approach by Gatti et al. [20] employs slogans and similar well-known expressions in generating interesting news titles. The generation process extracts keywords from a news article and then alters man-made slogans based on semantic similarities, dependency statistics, and other criteria, resulting in catchy news headlines.

Our proposed method for generating expressions differs from existing methods as follows. It focuses on generating slogans for a product while expressing a single adjectival property. We want the property to be expressed indirectly and metaphorically. Furthermore, our method creates slogans whilst considering one skeleton at a time, unlike the work carried in [48, 47] where several skeletons were considered. Producing metaphorical expressions is addressed in *Figure8*, which in contrast is concentrated on similes. Lastly, in our research, we examine the internal evaluation functions of our method against man-made slogans to highlight the impact of adopting them in generating slogans, an essential step missing from the current methods.

4 Collecting Slogans Corpora

We collect corpora of slogans, to get our hands on slogans produced by professionals for known advertising campaigns. Throughout this thesis, we use three corpora[8]. The first corpus, ψ_t contains slogans solely obtained from http://www.textart.ru/, due to its consistent structure of listing slogans and wide coverage of slogans from different categories. The corpus includes additional information regarding slogans such as the name of the brand and its category (e.g. pizza and university). In total, ψ_t has 3538 unique slogans. We mainly use ψ_t in our automatic analysis, to observe general linguistic characteristics in slogans. Additionally, we use it for evaluating certain aspects in man-made slogans in section 7.3.

As slogans in http://www.textart.ru/ might not be recent and famous, we crawled existing slogans from various online websites:

- http://www.namedevelopment.com
- http://www.taglineguru.com
- http://www.adslogans.co.uk
- http://www.advergize.com

As a result of the crawling process, we collected 6703 slogans. For each of these slogans, we store the name of the brand, category, year, product, media, country, and the source, when possible. As the retrieved slogans might be repeated, given that they were collected from different sources, slogans were given a fingerprint. The fingerprint was composed by stripping all white spaces and punctuation from the slogan and converting it to lower case. Slogans with the same fingerprint were grouped together. This process yields a total of 5824 unique slogans which constitute our entire corpora ψ_a. Random samples of slogans in ψ_a are used in our manual and semi-automatic analysis.

[8] Collected on 24 Oct 2016.

In addition to the above two corpora, we build a corpus, ψ_m, consisting of 40 well-known good and modern slogans[9]. Slogans in ψ_m are used as syntactic bases for generating slogans..

5 Manual Analysis of Slogans

We explain our manual analysis of existing slogans crafted by humans in this section. The main objective of analysing slogans manually is to highlight figurative characteristics in slogans. A similar analysis was performed in the literature by Miller and Toman [37]. However, we will analyse slogans using Mcquarrie and Mick's taxonomy, given in Figure 1, whereas the analysis by Miller and Toman followed a different taxonomy.

For this manual analysis, we randomly choose 100 slogans from ψ_a. We examine whether they contain figurative devices. In case the slogan contained at least one figurative device, we record their figurative modes (tropic or schematic), their rhetorical operations and, lastly, the type of figurative devices that were in them. This enables us to categorise slogans in the taxonomy proposed by Mcquarrie and Mick [36]. Note that only during our manual analysis, we follow the definitions of rhetorical devices such as chime and alliteration as declared by Mcquarrie and Mick [36], i.e. "keywords in a phrase begin with identical sounds or letters" and "three or more repetitions of constants.", respectively. Refer to Table 15 in the Appendix for the descriptions of the rhetorical devices covered in the taxonomy, as given by Mcquarrie and Mick.

The result of our manual analysis is that 91% of slogans contained at least one rhetorical device. During our annotation, we determined whether a rhetorical device existed by viewing it from different contexts. This is because slogans can be given without a context, and in case a context was given, it is typically the category and/or product. The high ratio of rhetorical devices in slogans shows that slogans are rich with figurative devices. Examples of slogans that did not contain any rhetorical devices are *Solitair*'s –a tour operator– slogan "Exclusive holidays for the single traveller." and *Two Men and a Truck*'s –a moving company– slogan "Movers who care.". In our analysis, literal slogans tend to describe the type of service, vision or product, or use a common expression. However, there are more cases of literal

[9] Retrieved from http://www.advergize.com/advertising/40-best-advertising-slogans-modern-brands/2/

slogans [36].

In the 91 figurative slogans, 37 and 33 out of them contained schematic and tropic figurative devices only, respectively. On the other hand, there were 21 slogans that had both figurative modes. "When it absolutely, positively has to be there overnight.", *FedEx*'s slogan, is an example of a slogan comprising schematic and tropic figurative mode. This slogan has two rhetorical operations, repetition and substitution, and two figurative devices, namely: alliteration and ellipsis. Another one is "Some cars fake it. These make it.". It is a slogan made by the automobile company *Chevrolet*. The slogan has the same rhetorical operations but different figurative devices, which are rhyme, parison, epistrophe, and ellipsis. A slogan that has a different tropic operation, destabilization, is *Volkswagen*'s slogan "Relieves Gas Pains". It has a metaphoric expression which indicates that their cars are seen as a healthy medicine that reduces the painful gases in the body. In addition to the use of a metaphor, it utilises alliteration. In these 21 slogans, only one of them contained reversal operations. It is *Volkswagen*'s slogan for their *GTI* cars, "Volkswagen GTI. For boys who were always men.". We can observe in the slogan the usage of the following rhetorical devices: antithesis, ellipsis, and paradox.

Regarding rhetorical operations, the most common rhetorical operation found in the 91 slogans is repetition, which appeared 58 times. The next frequent operation is destabilization with 34 occurrences followed substitution with 31 occurrences. The least used operation is the reversal as it only existed in three slogans.

Combinations of Rhetorical Operations	**Count**
Repetition (alone)	35
Destabilization (alone)	13
Substitution (alone)	11
Repetition, Destabilization	10
Repetition, Substitution	9
Substitution, Destabilization	8
Repetition, Reversal	2
Repetition, Substitution, Destabilization	2
Reversal, Substitution, Destabilization	1

Table 1: The number of times combinations of rhetorical operations appeared in slogans.

In Table 1, we show the count of found combinations of rhetorical operations in slogans. Looking at the stats, we observe that there was no slogan having all the four rhetorical operations. Furthermore, we notice that only three slogans had three different rhetorical operations. One of them is *Volkswagen*'s slogan for their *GTI* advertising campaign. One of the other two slogans is a slogan for an energy drinks company, *Golazo*. The slogan is "Golazo. Born to score.". In this slogan, assonance, ellipsis, and metaphor rhetorical devices were used. The metaphor in this slogan is in the form of personification, where an attribute of living creatures (being born) is assigned to the energy drink.

Moreover, we observe that the rhetorical operations repetition and destabilization co-occur together the most, 10 times. An example of such slogan is *Redwood Creek*'s –a wine company– "Satisfy your taste for Adventure!". This slogan demonstrates the usage of metaphor and alliteration.

The following frequent rhetorical operations occurring together are repetition and substitution. They have occurred 9 times. *KFC –Kentucky Fried Chicken–*'s slogan "Finger lickin' good." is an example of slogans containing repetition and substitution rhetorical operations. *KFC*'s slogan demonstrates the usage of assonance, ellipsis, and hyperbole.

Substitution and destabilization rhetorical operations co-occurred 8 times. 7 out of the 8 slogans had at least ellipsis and metaphor rhetorical devices in them, e.g. *Kraft*'s slogan "Kraft. Bite into summer". In addition to ellipsis and metaphor, the slogan of a bread company *–Burgen–* "Burgen. Harnessing the Power of Nature." has hyperbole in it. On the other hand, *Chrysler*'s slogan "Inspiration comes standard." contains only a personification metaphor and a hyperbole, i.e. no elliptical usages.

In our dataset, we notice that two slogans that used repetition and reversal rhetorical operations achieve that by using parison and antithesis, as seen in *Kix Cereal*'s slogan "Kid tested. Mother approved.".

Table 2 provides the frequencies of figurative devices in the 100 manually analysed slogans. We now analyse each figurative device individually. The goal of this analysis is to highlight any criteria to be defined in a computational function to detect or evaluate them. We focus more on the most frequent figurative devices.

Metaphor Metaphor is the most tropic figurative device used in slogans as it appeared in 27% of figurative slogans. Taking a closer look at these metaphorical slogans, we notice that the use of personification is common. Most of the time, human qualities are expressed using verbs as seen in *Glazo*'s slogan. Another example

Figurative Device	**Count**
Metaphor	27
Alliteration	26
Ellipsis	26
Assonance	16
Parison	15
Chime	14
Rhyme	9
Pun	5
Hyperbole	5
Anaphora	4
Metonym	4
Antithesis	3
Epistrophe	3
Paradox	2
Rhetorical question	2
Epanalepsis	1

Table 2: The frequency of each figurative device in slogans sorted from most frequent to least.

is “So good, it speaks for itself.”, *Blue Bunny*’s –an ice-cream company– slogan. A personification example that uses human adjectives is the slogan of a sewing machine, *Pffaf*, “Feel the creative energy.”. This slogan also holds another metaphor, which indicates that *energy* is a sensible object and can be felt. Another observation is that slogans use conventional metaphors, such as *Worcester Public Library*’s slogan “Your open door to opportunity.” and *Colorado Springs*’s “A company built on a reputation.”. Some slogans might convey a metaphor in a complex form such as *Diners Club*’s “The international symbol for yes.”. In this case, the entire expression can be seen as the vehicle of the metaphor, where the tenor is the company. On the other hand, *Nissan*’s slogan “Life is a Journey. Enjoy the Ride.” expresses the metaphor in a simple form, i.e. “*tenor* is a *vehicle*”, in addition to using a widely known metaphor. Another metaphorical expression is *Visa*’s “Life flows better with Visa.”. Here, life is visualised as a continuously flowing river. Nouns can also be used to stimulate metaphoric visualisations, as found in “A grip on the future.” and in “Your health in bloom.”, by *Bridgestone tires* and *Marin Acupuncture Clinic in*

San Anselmo, respectively.

From our metaphor analysis, we see that metaphors are expressed differently in slogans. They can be expressed in a simple form (e.g. a nominal metaphor) as seen in *Nissan*'s slogan, or they can be expressed in a very sophisticated fashion as in *Diners Club*'s slogan. However, such extreme cases occurred only once in our dataset. Personification was observed in multiple slogans, which was provoked by employing human-like attributes as adjectives and verbs. Adjectives and verbs seem to be used more frequently to express a metaphor than nouns, regardless of it being personification or not.

Ellipsis Ellipsis figurative device was used in many slogans. Twelve of the elliptic slogans started by the brand or product name followed by a full stop and incomplete sentence. In most of these elliptic cases, the verb to be is missing. "Volkswagen GTI. For boys who were always men." and "Lee. The jeans that built America." are examples of such cases. However, more might be omitted and provided in the context instead. For example, the slogan "Ford. Bold moves." was advertised[10] to indicate that "Ford *is for people who have* bold moves.". The verb to be is also be removed, as in *KFC*'s slogan "Finger lickin' good.". A common case of linguistic ellipsis is pseudogapping where part of the verb is omitted. The slogan "Service you expect at a price you wouldn't." by *Charles Hurst*, a car dealer company, employs pseudogapping where the verb *expect* is omitted from the end of the phrase. "Where you'd rather be!" and "When it absolutely, positively has to be there overnight." are subordinate clauses having the main clause missing. The main clause in these cases could be "*Product* offers/provides/..." or "Use *Product* ". These analyses suggest that it is possible to predict whether a slogan is elliptic or not by examining whether the slogan falls under any of the found scenarios, e.g. does it end with an auxiliary verb.

Rhyme In our rhyme analysis, we observe that a slogan might have more than two rhyming words. Some might have a rhyming effect between the last words of two phrases, e.g. "Some cars fake it. These make it." where two-word chunks rhyme at the end of the phrases. In some slogans, the rhyming words are close to each other, such as *Bissell*'s "We mean clean.", *Eric Brantner Copywriting*'s "Turning conversations into conversions." and *Applewood*'s "It's all good at Applewood.".

[10] YouTube copy of the ad: https://www.youtube.com/watch?v=Lc83WrZymJo

The slogan by *Mother Dairy*, “Mother Dairy. The creamier, tastier butter.”, has more than two words rhyming. In rare cases, all words in the slogan rhyme with each other, e.g. “Bigger! Better!” by *WindMill*. These variations suggest that rhyming effect can be found anywhere within the slogan.

Assonance Regarding assonance, we notice that they can be placed anywhere. Some of the rhyming examples used assonance such as *Bissell*’s slogan. *Miller-Coors*’s “It won’t slow you down” and *Casamigos*’s “Brought to you by those who drink it.” are additional examples of slogans employing assonance.

Chime The case with chime is similar to rhyme and assonance, pointing out that these prosody functions are employed in the same manner. Examples of chime usage include *Eric Brantner Copywriting*’s slogan “Turning conversations into conversions.” and *Belmont Barbershop*’s slogan “Clean cuts. Close shaves.”.

Alliteration Alliteration is no different from assonance, in terms of its usage and analysis. We notice that the constant t is used frequently to highlight alliteration, 14 cases out of 16 were found. For instance, *Rosella*’s slogan repeats it three constitutive times, *Tim Hortons*’s “It’s time for Tims.” three times and *Sole Clinic*’s “You trusted foot & rehab specialist.” four times. Other examples include the repetition of d as seen in *Kix Cereal*’s slogans, and of l as in *Kingsmill*’s slogan “Love bread. Love Kingsmill.”.

Parison Parison is one of the commonly used figurative devices in slogans. It is a rhetorical device on the syntactical level. *Belmont Barbershop*’s slogan (“Clean cuts. Close shaves.”) has parison in it as the structure of the phrases is repeated. In our analysis, we found that it is not a requirement for the structure to be identically repeated. For instance, *Goodyear Tires*’s slogan “Goodyear. We discover, you explore.” contains three phrases but only two of them have parison. As a result, as long as the similarities between the syntactical trees of the phrases in the slogan are high, we should consider it having pairson figurative device.

Hyperbole Few slogans, only 5, had hyperbole in them. For instance, the phrase “Harnessing the Power of Nature” is typically used in language to suggest that some massive natural resource is being utilised as a power source (e.g. waterfalls, wind, nuclear and sunshine). Using this expression in a slogan for a bread company, *Burgen*, is an exaggeration, in our opinion. "Live in water." by *Dolphin Swimwear* is another slogan that has exaggeration. However, its exaggeration is on the level of what the product offers not what the company does. A different form of hyperbole is *Sun Microsystems*' "The Network Is The Computer". In this understatement, the entire network (which connects multiple computers) is denoted as a singular computer. Similarly the case with "Inspiration comes standard." as there is no standardised form of inspiration.

Metonym Metonym is found in four slogans, namely: i) *Ford*'s "Ford. Bold moves.", ii) *CASA of Southwest Missouri*'s "Stand up for a child", iii) *Burger King*'s "Wake up with the King", and iv) *Sun Microsystems*' "The Network Is The Computer". In *Ford*'s and *Burger King*'s, the company name or part of it is used to represent the different products they have. On the other hand, *Sun Microsystems*' and *CASA*'s slogans use a singular form of the noun to refer to the whole. For example, "a child" in "Stand up for a child" refers to all abused and neglected children.

Anaphora, Epistrophe, Epanalepsis and Anadiplosis Some slogans have repeated words. Such repetition emphasises the importance of these words, plus it adds phonetic repetitions. Depending on how the words are repeated, they are considered as anaphora, epistrophe, epanalepsis and anadiplosis. Three out of the four rhetorical devices were found during our analysis, specifically: anaphora, epistrophe, and epanalepsis. Examples of anaphora are *Kingsmill*'s "Love bread. Love Kingsmill." and *Colour Line Hair Studio*'s "Love your hair, love yourself.". *Brentwood Chiropractic Clinic*'s and *Chevrolet*'s slogans "Feel better. Move better. Live better." and "Some cars fake it. These make it.", respectively, demonstrate epistrophe usages. "I am what I am." by *Reebok* employs epanalepsis. However, in *Reebok*'s slogan, it is a compound word that gets repeated, "I am".

Antithesis *Volkswagen*'s slogan: "Volkswagen GTI. For boys who were always men." along with other two utilise antithesis, where the opposite meaning of a word is also used in it. In this slogan, an opposite of the word "boys" was used, "men". The other examples are *Huggies Pull-Ups Training Pants*'s "End of nappies, start of pants." and *Kix Cereal*'s "Kid tested. Mother approved.".

Paradox Paradox was found in two slogans only, *Volkswagen GTI*'s and *Kraft*'s. In *Volkswagen*'s slogan, "Volkswagen GTI. For boys who were always men.", technically, boys cannot be men at the same time by definition. On the other hand, *Kraft*'s slogan, "And America spells cheese...K-R-A-F-T." is false by the definition of spelling words; therefore, considered paradox.

Rhetorical question Slogans are not expected to have responses. Hence, any question found in slogans is considered as a rhetorical question. "What are you up for?" by *Boyne Mountain Resort* and "Wouldn't you rather be Hemeling?" by *Hemeling* are the two slogans observed with rhetorical questions.

Pun Five puns were found in the analysed data. The word *King* in *Burger King*'s slogan "Wake up with the King" illustrates a pun as the word can be interpreted as a human king or as the brand's name. The brand name *Kingsmill*, also, can have two interpretations. In the company's slogan "Love bread. Love Kingsmill.", *Kingsmill* might be viewed as the brand's name or as *king's mill*. An example of a pun, where the pun is not on the brand name, is *US Mortgage Corporation*'s slogan "Helping you make it home.". In this slogan, the word *home* can mean a real estate or a comfortable place to live in.

6 Semi-Automatic Analysis of Slogans

Following our manual analysis, we perform a semi-automatic analysis on the manually annotated slogans. In this analysis, we analyse the slogans using multiple computational tools and services, and, then, we manually annotate their output to evaluate them. The purpose of this analysis is to investigate which computational resources are the most suitable for analysing slogans and short expressions prior to the fully-automated analysis.

Here, we automatically analyse the 100 randomly selected slogans from different points of view. Which are: (1) category prediction, (2) sentiment, (3) syntactical parsing and (4) entity recognition. Hereafter, we manually annotate the results obtained from the automatic analysis to assay their quality.

Category prediction In category prediction, we aim to discover whether the keywords used in slogans are strongly related to the category of the product. In our analysis, we use *Interactive Advertising Bureau (IAB) Quality Assurance Guidelines (QAG) Taxonomy*[11]. *IAB QAG* taxonomy provides a hierarchical classification of categories of advertised products and services, where the first level is abstract such as *Business, Arts* and *Food*, and the second level is a detailed classification, e.g. *advertising, books and vegan food.* For our analysis, we employed an online web service for text analysis called *Aylien*[12]. More specifically, their *IAB-QAG* classification service. *IAB*'s taxonomies get updated to reflect changes in the industry. To our best knowledge, the version of taxonomies provided by *Aylien* and used in our evaluation is 1.5. The classifier was able to provide predictions for 62 slogans with high confidence, and 83 slogans with low confidence. Overall, at least one prediction was produced for each slogan. An example of predictions retrieved for the slogan "Love bread. Love Kingsmill." are *Desserts & Baking*, and *Hobbies & Interests* with high confidence and *Hate Content, Non-Standard Content, Food & Drink,* and *Scrapbooking* with low confidence.

To examine the quality of the automatic classification, we manually annotated the classification results by determining whether the category of the company was correctly predicted (with high or low confidence) by the classifier. Based on our annotations, we observed that the classifier has successfully predicted the category of the company based on the text in the slogan for only 28 slogans. On the other hand, the category of 32 slogans was also correctly predicted but with low confidence. Nevertheless, only 12 slogans had both, low and high, correct predictions. We noticed that the classifier was able to suggest correct predictions in the following cases:

- The slogan contained the brand name and the brand is very well known (e.g. "Volkswagen GTI. For boys who were always men.")
- The slogan contained keywords that are strongly related to the category (e.g. *Sun Microsystems*'s "The Network Is The Computer" and *Marin Acupuncture Clinic*'s "Your health in bloom.")
- In rare cases, the slogan is very well known that you can predict the brand by the slogan (e.g. *PlayStation*'s "Live in your world. Play in ours.")

Given the low accuracy and our observations, we believe that employing such analysis in our generation model would not be advantageous.

[11] https://web.archive.org/web/20150912043429/http://www.iab.net/QAGInitiative/overview/taxonomy

[12] https://aylien.com/

Sentiment To analyse the sentiment of slogans, we experiment with two tools. The first is using *Aylien*'s text analysis API to predict whether a given sentence has a positive or negative sentiment. The other tool is an NLP library implemented in Python called Pattern [12]. The reasons for experimenting with these two tools are that *Aylien* is a web service continuously updated and trained to reach a higher accuracy of predictions; however, its free version is limited to 1,000 API requests per day. Pattern, on the other hand, is a free and open source library that offers a pre-trained model for predicting the sentiment of expressions. Comparing these models would give insights on how well the pre-trained model is.

The result of *Aylien*'s predictions is that 93% of slogans had positive sentiment. Nevertheless, some negatively predicted slogans are not necessarily negative, such as the slogans "Service you expect at a price you wouldn't." and "It won't slow you down.". Regarding the results obtained by *Pattern*, 95 slogans were considered positive with threshold ≥ 0.0, i.e. neutral slogans were treated as positive. Dividing neutral and positive slogans apart results in classifying 56 and 39 into the two categories, respectively. The predictions between the two tools were not always consistent. This means that some slogans were classified as positive by one tool and negative by the other. For instance, both tools classified "It won't slow you down" as negative; however, they did not agree on the same classifications for "Some cars fake it. These make it." and "Where you'd rather be!".

To evaluate the two tools, we manually annotate the sentiment of the slogans and compare the annotations to the predictions of the tools. From our annotations, it appears that all the slogans in the test set had a positive or neutral sentiment. This makes the test set one-sided as it only contains slogans with positive sentiment. Based on these analyses, we choose to use *Pattern* as the sentiment classifier in our method because it had the highest accuracy and it is free.

Syntax To analyse the syntax of slogans, we parse their dependencies using natural language processing tools. There exist multiple tools for performing the task; however, the quality of parsing varies. As a result, we evaluate four of the most well known and accurate tools for the task, namely *Stanford CoreNLP* [33], *Pattern* [12], *Spacy* [26], and *SyntaxNet*[13] by using them to parse the hundred randomly selected slogans. There exist other syntax parsers that we did not consider in our analysis, such as *maltparser* [38], due to the scope of the thesis. In our analysis, we

[13] https://github.com/tensorflow/models/tree/master/research/syntaxnet

concentrate on comparing two parameters, accuracy and speed.

Slogans are typically non-formal short sentences. Given that most of the current natural language parsers are built and trained to process complete sentences, such in news, books, and web, in contrast to slogans, we need to choose a parser that handles incomplete sentences to a great degree. To measure the accuracy of the parser's result, we manually annotate whether its output is correct or not.

On the other hand, the speed of the parsing processes is an essential parameter in our case as we will be processing a huge amount of text; therefore, it enhances the efficiency of the system.

The results of the analysis is given in table 3. *Spacy* had the highest accuracy, 88%, of all the parsers followed by *Stanford CoreNLP* with 82% accuracy. *Pattern* achieved 66% accuracy whereas *SyntaxNet* had the least accuracy, 41 correct predictions out of 100. We believe that the reason *SytaxNet* had such low accuracy is the nature of slogans being incomplete.

Regarding the speed of the parsers, *Pattern* took the least time to parse the slogans, 0.7 seconds. The second top parsing candidate in regards to speed is *Spacy*, which completed the task in around 1.9 seconds. *SyntaxNet* and *Stanford CoreNLP* spent 2.6 and 6.9 seconds, in the same order.

An example of a slogan that all parsers failed to parse correctly is *Kingsmill*'s "Love bread. Love Kingsmill.". The parsers miss-predicted the part-of-speech of the word "Love" in either of the sentences as a noun or adjective. As a result, none of the parsers tagged the word "Love" as a verb.

Comparing the parsers, we notice that *Spacy* is the best candidate as it has the highest accuracy and, yet, efficient. Henceforth, we will be using *Spacy* as the natural language processing tool in this thesis.

Parser	**Accuracy**	**Total Time (s)**
Spacy	88	1.85
StanfordCoreNLP	82	6.87
Pattern	66	0.67
SyntaxNet	41	2.59

Table 3: The accuracy and total processing time the 100 randomly selected slogans.

Entity Detection Entity detection refers to identifying whether a given word in the text refers to an entity and recognising its type (e.g. organisation, person, number ...etc). Slogans might contain entities such as brand or product's name. We have manually examined the existence of such entities in the randomly selected slogans. We found that 3 in every 10 slogans contained a named entity, either the name of the brand or the product.

Furthermore, to automate such detection, we employ the named entity recognition feature in *Spacy*. We only consider entity names recognised as a person or organisation types because we believe other entity types (e.g. country names) are generally irrelevant for slogans. Comparing the manual analysis with *Spacy*'s prediction, we see that named entities existing in 12 slogans were correctly predicted. In some cases, false positive, predicting that a given word is an entity while it is not, has occurred, as seen in *Redwood Creek*'s slogan "Satisfy your taste for Adventure!" where Adventure was predicted as the organisation's name.

7 Computational Analysis and Generation of Slogans

In this section, we explain our computational method for both analysing and generating slogans.

The goal of this thesis is to propose a computational method for generating slogans. To generate natural language, we need to employ computational natural language resources (e.g. corpora and language models). We start by describing the resources used in our method, in section 7.1. Thereafter, we define a list of features for measuring and estimating some linguistic characteristics, based on our manual analysis and related work. Examples of such features are prosody (sound repetition rhetorical devices such as assonance and rhyme) and metaphors.

As per our manual analysis and the research conducted by Miller and Toman [37], slogans tend to be metaphoric. Hence, we propose a method for generating apt metaphors for highlighting the input adjectival property in the concept. Our proposed method utilises a metaphor interpretation model. The proposed method is a key contribution of this thesis.

Following our metaphor generation method, we delineate our slogan generation method in section 7.5. Our slogan generation method uses the generated metaphors along with the computational resources to produce slogan candidates.

To facilitate referencing the notations defined in this thesis, we list all of them under the Appendix in Table 16.

7.1 Computational & Linguistic Resources

This section covers the linguistic and computational resources used in the proposed method.

Text Corpus, ζ: We use a 2 billion word web-based text corpus, *ukWaC*[14], as the main corpus. All corpus-based models in our approach are built using this corpus. We chose a web-based corpus to cover a wide range of topics and different writing styles.

Language model, ξ: We build a probabilistic bigram language model ξ using bigram frequencies provided with *ukWaC*. The language model (ξ) is built to estimate the probability of a created slogan by our method to be generated it. A slogan with high probability is more likely to be grammatically correct as it appeared frequently in the corpus ζ. Employing bigrams, in contrast to trigrams or higher n-grams, gives the method a greater degree of freedom in its generations. Higher n-grams would improve the grammar of the generated expressions but would tie them to expressions in the original corpus.

Semantic model, ω: The goal of constructing a semantic model is to find words that are semantically related to another word and to measure the semantic relatedness between two words. We follow the approach described in *Meta4meaning* [53] in building the semantic model ω. In order to do that, we start by obtaining co-occurrence counts of words in ζ, constrained by sentence boundaries, within a window of ± 4. We limit the vocabulary of the model to the most frequent 50,000 words, excluding closed class words. We then convert co-occurrence counts to relatedness measure by employing the log-likelihood measure defined by [16] while capping all negative values to zero. Finally, we normalize relatedness scores using L1-norm [35].

[14] http://wacky.sslmit.unibo.it

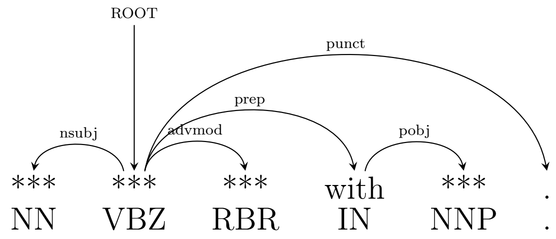


Figure 2: An example of a skeleton constructed from Visa's slogan: "Life flows better with Visa."

Expression skeletons, δ: A slogan skeleton is a parse tree of a sentence where all content words are replaced with a placeholder "***", i.e. stop words are kept. Nevertheless, all grammatical relations between words and part-of-speech tags are maintained. The purpose of using a database of skeletons is to reuse syntactical structures of effective slogans. By observing some well-known slogans, we notice that slogans might share syntactical structures, e.g. *Volkswagen*'s "Think Small." and *Apple*'s "Think Different.".

We utilise *Spacy*[15] as a natural language processing tool to parse ψ_m (the 40 well-known slogans). Prior to constructing the skeletons, we preprocess the slogans to increase the parsing accuracy. The first preprocessing step is converting capitalized words into lower case, except the first word and any recognized named entities. This step reduces miss-classifying verbs, adverbs and adjectives as nouns (e.g. the adverb *differently* in *Red Lobster*'s slogan "Seafood Differently."), yet they could occur as many slogans are not complete sentences. Slogans tend to be informal; therefore, we convert words with the suffix *VERB*-in' into *VERB*-ing, in the second step. As a result of the preprocessing phase, *KFC*'s slogan "Finger Lickin' Good." becomes "Finger licking good.".

Subsequently, we convert slogans into skeletons. Figure 2 provides an example of a skeleton generated from *Visa*'s slogan "Life flows better with Visa.".

Once all slogans are transformed into skeletons, we only keep skeletons that have at least 40% of their tokens as placeholders and have a minimum of two placeholders. These conditions ensure that the method has some freedom in filling in the skeleton. As a result, slogans such as *Reebok*'s "I am what I am." and *Coca-Cola*'s "Enjoy." are removed[16]. In total, the database contained 26 unique skeletons.

[15] http://www.spacy.io

[16] Because their skeletons "I am what I am ." and "***_NOUN .", respectively, are very restrictive. The first has no placeholders to fill while the second has only one.

We also store additional meta-information in skeletons such as entities recognised and their type (e.g. product, organization, person ...etc) to assess in substituting entity names with a desired product name.

Grammatical relations, γ: A grammatical relation is a single relation in the parse tree which contains a word (called dependent), its head word (called governor), the parts-of-speech (based on Penn Treebank tag set) of both words, and the type of relation. Similarly to approaches by Özbal, Pighin, and Strapparava [39] and Tomašič, Żnidaršič, and Papa [48], we build a repository of grammatical relations. We parse the entire corpus ζ using *Spacy* and store all grammatical relations observed along with their frequencies. We retain grammatical relations with frequencies ≥ 50 to remove rare and noisy cases. The process yields 3,178,649 grammatical relations[17].

Grammatical relations are very useful in general and creative NLG tasks. For instance, Hämäläinen [24] have used similarly constructed grammatical relations of Finnish [23] in creating Finnish poetry. Despite the significant importance of such relations, there are no publicly available repositories of English grammatical relations, to the best of our knowledge. Therefore, we release the repository of grammatical relations [4], in addition to another repository containing grammatical relations with frequencies ≥ 10. These repositories contribute to the available resources for NLP and NLG.

Nouns and Their Adjectival Properties, κ We employ two resources for retrieving nouns associated with the input property. The first resource, $\kappa_{General}$, is *Thesaurus Rex* [51]. *Thesaurus Rex* is used for retrieving general nouns (e.g. coffee, flower, ...etc). On the other hand, the resource provided by Alnajjar et al. [2], κ_{Human}, is employed to obtain nouns of human categories (e.g. actor, lawyer, politician, ...etc). These resources will be used in generating metaphors, the former for general metaphors and the latter for personifications.

[17] To obtain and count the frequencies of these relations, we have utilised a computing cluster provided by the Department of Computer Science at University of Helsinki, *Ukko2*. To process the entire corpus efficiently, we have employed a MapReduce approach where the entire task is divided and assigned to multiple computing resources (7 nodes with 40 CPUs) and, once done, their outputs were reduced and merged into a final single output. The total time for completing the task was 21.4 hours.

7.2 Functions for Measuring Linguistic Characteristics

To be able to generate slogans containing certain linguistic aesthetics (e.g. linguistic devices and semantics), we define functions that are intended to measure such aesthetics. Our hypothesis is that the criteria that we define below are important for generating slogans. We test this hypothesis by 1) observing the usage of these functions in existing slogans during our automatic analysis and 2) employing them in our slogan generation method and, then, evaluating the slogans output by it.

Functions for measuring a given objective are denoted as $f_{function\_name}$, where *function_name* is the name of the function. As an example, $f_{consonance}$ would be a function for measuring the usage of consonance in a slogan.

The following are the characteristics considered by us: (1) number of tokens, (2) sentiments, (3) semantic relatedness, (4) word infrequencies, (5) grammatical relations, (6) prosody and (7) metaphoricity. From the manual (section 5) and semi-automatic analysis (section 6), we notice that the sentiment, prosody and metaphoricity are common characteristics of human-made slogans; therefore, we define computational functions for measuring them. Inspired by Özbal, Pighin, and Strapparava [39], we use the unusual-words scorer function to estimate the surprisingness effect caused by infrequent words in slogans. Regarding semantic relatedness and grammatical relations, we believe that they are essential characteristics for suitable slogans. We look at the number of tokens that man-made slogans have solely to reveal statistics on what is the minimum, maximum and average length of slogans, i.e. the function is not used in the method for generating slogans.

This section describes our implementations of computational functions for measuring the above linguistic characteristics in slogans.

Throughout this thesis, we use the following notations as introduced in section 2.3. T and P represent the target concept and an adjectival property to highlight in the target concept, respectively. $\mathcal{E}$ denotes an expression (slogan). A slogan $\mathcal{E}$ is a sequence of tokens $\{t_1, t_2, t_3 \dots t_m\}, m \geq 1$. Content words in a slogan are indicated as $c(\mathcal{E}) = \{\mathcal{E} \setminus stop\}$, where *stop* is the set of stop words. Given these notations, we define the functions as below.

Number of Tokens (f_{len}): As slogans are short expressions, we measure the number of tokens in slogans to know their acceptable range and average length.

In short, f_{len} measures the number of tokens in a given expression, which is defined as

$$f_{len}(\mathcal{E}) = |\mathcal{E}|. \tag{1}$$

Sentiment ($f_{positive}$): Advertising slogans tend to contain positive words [14] which would give the receiver a positive feeling about the brand. Also, our semi-automatic analyses, in section 6, have shown that automatic sentiment prediction models have detected positive sentiment in most slogans. As a result, it is important to employ sentiment analysis in producing slogans. We employ a sentiment classifier provided in *Pattern* [12] to predict the sentiment polarity score of sentences. Sentiment polarity score is a value between -1.0 and +1.0. Denoting the classifier as *sentiment*, we define the sentiment feature as follows:

$$f_{positive}(\mathcal{E}) = \begin{cases} 1, & \text{if } \mathit{sentiment}(\mathcal{E}) \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

We consider slogans with a neutral sentiment, i.e. $\mathit{sentiment}(\mathcal{E}) = 0$, as positive expressions because slogans are short by nature and might not necessarily contain positive words, and to reduce miss-classifications.

Semantic Relatedness (f_{rel}): The semantic relatedness function accepts two parameters, the expression $\mathcal{E}$ and a target word η. Using the semantic model ω, the function then computes the mean of semantic relatedness between all content words and the target word as:

$$f_{rel}(\mathcal{E}, \eta) = \frac{\sum_{t_i \in c(\mathcal{E})} \omega(t_i, \eta)}{|c(\mathcal{E})|} \tag{3}$$

Semantic Cohesion ($f_{cohesion}$): Semantic cohesion is performed to ensure that all words in the expression are semantically related to each other.

$$f_{cohesion}(\mathcal{E}) = \begin{cases} 1, & \text{if } \omega(t_i, t_j) > 0,\ \forall_{t_i}, \forall_{t_j} \in c(\mathcal{E}) \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Prosody ($f_{prosody}$): In our work, we consider four features of prosody, namely: i) rhyme, ii) alliteration, iii) assonance and iv) consonance. We make use of *The CMU Pronouncing Dictionary* [31] to measure the frequency of repeated sounds in words.

Let $\varphi(t)$ be *CMU*'s function which returns phonemes in a given word t, where $\varphi(t) = (p_1, p_2, \ldots, p_t)$. Equation 5 is for counting the total number of occurrences of a phoneme, *pho*, in an expression $\mathcal{E}$.

$$count_{phoneme}(\mathcal{E}, pho) = \sum_{t \in \mathcal{E}} |\{i \mid \varphi(t)_i = pho\}| \tag{5}$$

Following the definitions by Mcquarrie and Mick [36], we implement the features for measuring assonance and consonance in slogans. Let *vowels* be a set containing phonetic transcriptions of vowels. We count the number of vowels and constants existing in a given expression as in the equations 6a and 6b, in the given order. As per the definitions [36], we only consider sounds repeated at least three times, i.e. $count_{phoneme}(\mathcal{E}, pho) \geq 3$.

$$count_{assonance}(\mathcal{E}) = \sum_{pho \in vowels} count_{phoneme}(\mathcal{E}, pho) \tag{6a}$$

$$count_{consonance}(\mathcal{E}) = \sum_{pho \notin vowels} count_{phoneme}(\mathcal{E}, pho) \tag{6b}$$

Using the above equations, we implement the assonance and consonance functions as below.

$$f_{assonance}(\mathcal{E}) = \frac{count_{assonance}(\mathcal{E})}{|\varphi(\mathcal{E})|} \tag{7a}$$

$$f_{consonance}(\mathcal{E}) = \frac{count_{assonance}(\mathcal{E})}{|\varphi(\mathcal{E})|} \tag{7b}$$

Regarding alliteration and rhyme, the functions count the number of tokens having matching phonemes, at the beginning and end of the expressions, respectively. In our alliteration and rhyme implementations, we prefer quantity over quality, i.e. find matching sounds between words regardless of the quality and stress. For simplicity, we only consider phonemes; hence, syllables are not taken into account. Both functions utilise an extra function, *matching_pho*, which accepts two phonemes, in our case, and returns 1 if they matched and 0 otherwise.

$$matching\_pho(a, b) = \begin{cases} 1, & \text{if } a = b \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

We denote by $\varphi(t)_0$ the first phoneme and by $\varphi(t)_{-1}$ the last phoneme in a word t. We measure alliteration and rhyme as follows:

$$f_{alliteration}(\mathcal{E}) = \frac{\sum_{t_i} \sum_{t_{j \neq i}} matching\_pho(\varphi(t_i)_0, \varphi(t_j)_0)}{|\mathcal{E}|} \tag{9a}$$

$$f_{rhyme}(\mathcal{E}) = \frac{\sum_{t_i} \sum_{t_{j \neq i}} matching\_pho(\varphi(t_i)_{-1}, \varphi(t_j)_{-1})}{|\mathcal{E}|} \tag{9b}$$

Word Infrequencies ($f_{unusual}$): Inspired by Özbal, Pighin, and Strapparava [39], we define a function for estimating surprisingness. The function measures how infrequent, i.e. unusual, the individual words in the slogan are, which is implemented as

$$f_{unusual}(\mathcal{E}) = \frac{\sum_{t \in c(\mathcal{E})} \frac{1}{freq(t)}}{|c(\mathcal{E})|}, \tag{10}$$

where $freq(t)$ indicates the number of times the word t was observed in the corpus ζ. In case all content words did not appear in the corpus, 0 is returned as the surprisingness score.

Metaphoricity ($f_{metaphoricity}$): The metaphoricity function contains two sub-functions. In these functions, we assume that an apt metaphorical vehicle v for tenor T is given. The first function aims at measuring how the content words $c(\mathcal{E})$ in the slogan $\mathcal{E}$ are related to both, the tenor T and the vehicle v. This relatedness feature and the first metaphoricity function are measured as follows:

$$max\_rel(\mathcal{E}, x) = \operatorname*{argmax}_{t \in c(\mathcal{E})} \omega(t, x) \tag{11a}$$

$$f_{metaphoricity\_1}(\mathcal{E}, T, v) = max\_rel(\mathcal{E}, T) \cdot max\_rel(\mathcal{E}, v) \tag{11b}$$

$f_{metaphoricity\_1}$ is implemented as the product of the maximum relatedness value of words in an expression to the tenor and vehicle. By employing this function, we guarantee that the slogan contains a word that is related to the tenor and another that is related to the vehicle, when $f_{metaphoricity\_1} > 0$. The higher the value of the function, the more related the words are to the tenor and vehicle.

The other metaphoricity function is applied to support having at least one word that is strongly related to the metaphorical vehicle v but not to tenor T:

$$f_{metaphoricity\_2}(\mathcal{E}, T, v) = \underset{t \in c(\mathcal{E})}{\operatorname{argmax}}(\omega(t, v) - \omega(t, T)) \tag{12}$$

When interpreting an expression containing a word that is strongly related to the metaphorical vehicle v (but not to T) in the context of T, it is more likely to perceive it as metaphorical. For instance, let the tenor T be *car* and the vehicle v be *dancer*. The following expression $\mathcal{E} = (\text{“}cars\text{”}, \text{“}of\text{”}, \text{“}street\text{”}, \text{“.”})$ would contain a word (*street*) that is strongly related to the vehicle *dancer* (because of the street dancing style) but it is also strongly related to the tenor *car*. As a result, the expression would not be metaphorical. On the contrary, the word *stage* in the expression $\mathcal{E} = (\text{“}cars\text{”}, \text{“}of\text{”}, \text{“}stage\text{”}, \text{“.”})$ is strongly related to the vehicle but not to the tenor, which makes it metaphorical. $f_{metaphoricity\_2}$ is introduced to measure and encourage such metaphoricity.

Grammatical Relations ($f_{grammar}$): The last function checks whether all grammatical relations found in an expression exist in the grammatical relations repository γ. Let $\rho(\mathcal{E})$ refer to all grammatical relations in $\mathcal{E}$ returned by the natural language processing tool *Spacy*. Using ρ, we check grammatical relations as follows:

$$f_{grammar}(\mathcal{E}) = \begin{cases} 1, & \text{if } \rho(\mathcal{E}) \subset \gamma \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

7.3 Computational Analysis of Slogans

The purpose of our automatic analysis is to employ the defined computational functions, in section 7.2, in measuring how existing slogans hold the criteria which are evaluated by the functions. Such analysis aid in evaluating our definitions and discovering some criteria of good slogans. In this section, we provide our automatic analysis of the 3538 slogans collected in ψ_t on all the previously defined functions.

The first feature we examine is the length of slogans. Evaluating f_{len} on all slogans, we observed that the maximum number of tokens a slogan has is 18, including stop words and punctuation marks. The minimum is two tokens. An example of a slogan containing 18 tokens is “Get the care you need, when you need it, at the price you can afford.” by *Chiropractor Plus Clinic*. The shortest slogan is “Engage.” by *Furman University in South Carolina*. The average number of tokens is 6, such as in *XS Energy Drink*’s slogan “Premium energy. Explosive taste.”.

The second feature we analyse is the positive sentiment of slogans using $f_{positive}$. The vast majority, 95% of slogans, have positive sentiment. Some slogans were classified as negative expressions while they do not necessarily express any negative sentiment. An example of a slogan with a false negative classification is "It's hard to resist a Lu.", by a cookies producer *Lu Le Petit Ecolier*.

The next analysis we perform examines the semantic relatedness between content words and the product. Treating the meta-information regarding the category of the brand (e.g. university) of slogans collected in ψ_t as the target concept T, we evaluate the function $f_{rel}(\mathcal{E}, T)$ on all slogans. 7 in every 10 slogans had a relation to the target concept T, in other words $f_{rel}(\mathcal{E}, T) > 0$. For instance, *Granini*'s slogan "Granini. The fascination of fruit" had relatedness value of 0.09 to *juice*, the category of *Granini* in ψ_t. The remaining 30% of slogans did not contain words that are related to the category.

Using the semantic cohesion function, $f_{cohesion}$, we find how many slogans have their content words related to each other. In our test, only 15% slogans had semantic cohesion. This low number could be due to reasons such as content words not existing in the semantic model and the strict condition that every word has to be related to every other word. The total number of unique content words found in slogans in ψ_t is 4090. 40% of these words did not exist in the semantic model ω, e.g. infrequent entity names such as *Granini*. In terms of semantic cohesion between pairs of words, 76% of individual pairs were not related based on the semantic model.

Regarding prosody features, we measure sound repetitions using the four prosody functions, $f_{assonance}, f_{consonance}, f_{alliteration}, \text{and} f_{rhyme}$. We desire to investigate the number of occurrences of each prosody feature in slogans. Furthermore, we observe the overall number of different prosody features slogans have.

The most occurring prosody feature in our database was consonance, with 14% occurrences. This might be due to the fact that there are more consonants in English than vowels. *Puffs Tissues*'s slogan "A nose in need deserves Puffs indeed." is an example with consonance repetition of the consonants n and s.

The second most observed prosody feature is rhyme as it appeared in 13% of slogans. "Mother Dairy. The creamier, tastier butter.", by *Mother Dairy*, highlights rhyme between the words *Mother, creamier, tastier* and *butter*.

11% and 9% slogans contained alliterations and assonance, respectively. An example of a slogan utilising alliteration is "The property protection people." by *TNT Secu-*

rity Systems, whereas *Mother Dairy*'s slogan demonstrates the usage of assonance.

In summary, 17% of slogans contained at least one prosody feature. Approximately, half of them contained two or more prosody features. Three prosody features were found in 4% slogans, whereas only 20 slogans employed all four types of prosody. An example of such slogan is "The youngest and fastest fleet." by *Superfast Ferries*.

We next measure the surprisingness (i.e. word infrequencies) using $f_{unusual}$. The maximum surprisingness value a slogan had is 0.5, whereas the minimum was 0. On average, slogans had a surprisingness value of 0.004.

The last automated analysis we perform is checking whether all grammatical relations in a slogan appeared in the corpus ζ at least 50 times. Evaluating all the slogans using the function $f_{grammar}$, we noticed that all grammatical relations in all the 3538 slogans have appeared in the grammatical repository γ. This high coverage of relations might be due to the large size of the grammatical repository and the type of corpus.

7.4 Computational Generation of Metaphors

As per our analysis and related work in the literature, slogans seem to include metaphors, in addition to the positive effects using metaphors has on recalling the name of the brand. Accordingly, we propose a method for generating metaphors which will be used in our slogan generator to produce metaphorical slogans. The task of metaphor generation is defined in section 2.3. In short, the method should generate apt metaphorical vehicles that would highlight a certain property P in the target concept, tenor, T. An example of such input is T ="computer" and P ="creative".

For the input property P, the method begins by retrieving nouns associated with P using κ. We retrieve two types of nouns from the resource κ, general nouns from $\kappa_{General}$ and nouns of human categories from κ_{Human}. We use the top 10% of each type to only pick candidates strongly related to P.

The above procedure gives nouns related to the given property P, but it does not ensure that their metaphorical interpretation in the context of tenor T is P. To select nouns that are likely to have the intended interpretation, we employ a corpus-based metaphor interpretation model, *Meta4meaning* [53].

Meta4meaning accepts two nouns as input, a tenor and a vehicle, and produces a list of possible interpretations for the metaphor. To our knowledge, the proposed method here is the first for generating metaphors based on their interpretations.

Using *Meta4meaning*, the method interprets the potential metaphorical nouns retrieved by calculating the combined metaphor rank metric, c.f. Xiao et al. [53]. Only nouns with the property P among the top 50 interpretations are used. Additionally, as metaphors are asymmetrical, the method removes vehicle candidates that have the interpretation rank of "T is [a] v" greater than the interpretation of the reversed metaphor, i.e. "v is [a] T".

For example, nouns in κ that are strongly associated with $P =$ "*creative*" are:
$\kappa_{General}(creative) =$ {painting, music, ..., presentation}
$\kappa_{Human}(creative) =$ {artist, genius, poet, ..., dancer}

By interpreting these candidates using *Meta4meaning* and pruning out candidates not meeting the predefined conditions, we obtain the following candidates where the score is the interpretation rank (i.e. smaller is better):
$V_{General}(computer, creative) =$ {art: 4, drama: 4, director: 4, artist: 5, ..., exhibition: 50}
$V_{Human}(computer, creative) =$ {genius: 2, artist: 5, designer: 12, ..., inventor: 49}

Finally, we merge the two lists of potential vehicles into one, $V = V_{General} \cup V_{Human}$.

7.5 Computational Generation of Slogans

In this section we elaborate on the process of generating slogans. Our slogan generation method accepts three inputs, the target concept T, an adjectival property P to express and a metaphorical vehicle v which highlights P in T. Throughout this thesis, we will use vehicles generated from the metaphor generation process. Nonetheless, v can be input manually as well.

On a high-level, the slogan generation method reuses a syntactical structure (i.e. skeleton) learned from existing slogans as a base. To fill it with appropriate words that would result in a slogan, we employ an optimisation algorithm to produce expressions containing certain criteria (e.g. metaphoricity).

This section is divided as follows. We start by explaining how we construct the search space. Thereafter, we motivate and define the aspects which we will consider while finding potential solutions, followed by a detailed description of the generation algorithm.

Construction of Semantic Spaces From the pool of skeletons of well-known slogans δ (c.f. section 7.1), the method selects a skeleton s at random. Given a fixed skeleton s, the method then constructs two semantic spaces. In this work, a semantic space contains a set of words that could be used as potential fillers for s and have a semantic meaning in a context (i.e. exist in the semantic model ω). The two spaces that are constructed in the method are (1) interesting I and (2) universal Υ semantic spaces.

The interesting semantic space, which contains words that are favoured, is constructed by obtaining words, from ω, that are either related to the input concept T or the vehicle v. The method obtains the k words most strongly related to T. In our case, k was empirically set to 150. The method includes related words to v to encourage the generation of metaphorical expressions. Also, the top k related words to v, in ω, are collected while ensuring that they are abstract. This condition is applied because abstraction tends to be required in processing metaphors [22]. To select only abstract terms, we utilise the abstractness dataset provided by Turney et al. [49] and keep words with abstractness level ≥ 0.5. After all related words are obtained, we define I as:

$$I = \{a_i \in \omega(T) \mid i \leq 150\} \cup \{a_i \in \omega(v) \mid i \leq 150\} \tag{14}$$

We define Υ to be the total semantic space which contains all possible words that could fill s (i.e. have a part of speech and grammatical relation matching a placeholder in s, based on γ).

The search space of slogans, given a skeleton s, consists of all feasible ways of filling the skeleton with words in I or alternatively in Υ. The task of the expression generator is to traverse the search space and find suitable solutions (i.e. good slogans).

Criteria of Good Slogans We employ the functions defined in section 7.2 (except f_{len}) as criteria for good slogans. The function f_{len} is discarded because it has no effect on the generation process as generated slogans are based on existing slogans, i.e. they will have the same length. The selection of these criteria is part of the experimental work to evaluate them and their effects on generating good slogans.

We divide the criteria into two categories, filtering and evaluation. Filtering criteria exist to delete any expression that is not acceptable or invalid (boolean), whereas evaluation criteria are employed to be maximised (ratio).

In our method, the filtering criteria are i) relatedness between words within the slogan and ii) positive sentiment. On the other hand, the evaluation criteria consist of i) relatedness to the input T and P, ii) language correctness and word frequencies and iii) rhetorical devices. Depending on the overall creative goal, a different set of evaluation criteria could be used. For instance, to generate ironic expressions one might use negatively related terms.

Details of these criteria are explained in the remainder of this section, in the Filtering and Fitness Functions paragraphs.

Algorithm for Traversing the Search Space We employ genetic algorithms to find good slogans in the above-described space of possible slogans, given a fixed skeleton. We use Deap [17] as the evolutionary computation framework. Below, we use μ to denote the size of the population, G the number of generations to produce, and $Prob_m$ and $Prob_c$ the probability of the mutation and crossover, respectively.

Our algorithm first produces an initial population and then evolves it over a certain number of iterations. Starting with the initial population, the employed $(\mu + \lambda)$ evolutionary algorithm produces λ number of offspring by performing multiple crossovers and mutations. The algorithm then puts the current population and offspring through a filtering process (discussed below). The population for the next generation is produced by evaluating the current population and the offspring, and then selecting μ number of individuals. The evolutionary process ends after the specified number of generations.

Initial Population Given a skeleton s, our algorithm begins filling the word (slot) with the most dependent words to it, usually it is the root. Using the grammatical relations resource γ, the algorithm ensures that the words satisfy the grammatical relations of s. The algorithm attempts to randomly pick a word residing at the intersection of I and Υ, i.e. interesting and possible. If the intersection is empty, a word is randomly picked from the set of possible fillers Υ. The algorithm repeats the same process for filling the remainder of the words, also taking into account the conditions imposed by the already filled words. However, if the process fails to locate a suitable filler for the next word slot, the individual is discarded and the process starts over. The process continues until the desired number of individual expressions are generated, serving as the initial population.

Given the large knowledge bases used, especially the grammatical relations γ and semantic relatedness ω, it is unlikely for the approach to fail in creating slogans for a given input; however, it is yet possible in some cases such as (1) a rare concept or property with few or noisy associations, (2) a low k threshold or (3) a grammatically incorrect skeleton.

Mutation and Crossover Our algorithm employs only one kind of mutation. The mutation randomly selects and substitutes a word from the expression. In doing so, it follows the same process as was described for the slogan generation for the initial population. Our algorithm applies a one-point crossover. The resultant newly generated child expressions are then put through a grammatical check to verify that all grammatical relations in the expressions exist in our grammatical relations repository γ. A failure of the grammatical check, for any of the two children, results in their disposal while parent expressions are kept in the population.

Filtering The function $f_{cohesion}$ is used to check relatedness of words in the slogan against each other. The slogans with unrelated words are filtered out.

The filtering process also removes any expressions with negative sentiments, i.e. $f_{positive}(\mathcal{E}) \neq 1$.

The mutation and crossover may produce duplicate slogans. The filtering stage also takes care of such anomalies. Once a new generation is produced, the filtering process removes any duplicates.

Fitness Functions Based on the criteria of good slogans stated earlier, we define the internal evaluation metric of the genetic algorithm. We define four main dimensions to optimise: i) target relatedness, ii) language correctness, iii) metaphoricity and iv) prosody. Each dimension can be further composed of multiple sub-features. In the genetic algorithms, these sub-features are weighted and summed to represent the entire dimension. We have empirically assigned the weights of these functions.

Target relatedness measures the relatedness of the words in the slogan to the target input, i.e. T and P, using f_{rel}. $f_{rel}(\mathcal{E}, T)$ and $f_{rel}(\mathcal{E}, P)$ are two sub-features of the relatedness dimension.

The language dimension is concerned with how probable is the slogan to be generated with language model ξ. Additionally, surprisingness, $f_{unusual}(\mathcal{E})$, is another feature in the language dimension.

The metaphoricity dimension contains two metaphoricity functions ($f_{metaphoricity\_1}$ and $f_{metaphoricity\_2}$) which are explained in section 7.2.

Lastly, the fourth dimension covers the four features of prosody defined earlier, i.e. i) assonance ($f_{assonance}$), ii) consonance ($f_{consonance}$), iii) alliteration ($f_{alliteration}$) and iv) rhyme (f_{rhyme}).

Selection Some of the evaluations involved in our algorithm are conflicting in nature. An example of two conflicting dimensions is the relatedness and metaphoricity dimensions. This is because the relatedness dimension would maximise the relatedness of words in the slogan to input (i.e. target concept and property), while the metaphoricity dimension would maximise their relatedness to the metaphorical vehicle. A single sorting method for selection, based e.g. on the sum of all evaluations, could potentially lead to the dominance of one of the evaluations over others, resulting in imbalanced slogans. Therefore, our selection process involves a non-dominant sorting algorithm (NSGA-II) which is more effective when dealing with multiple conflicting objectives [13].

8 Empirical Evaluation

8.1 Evaluation Setup

We perform three empirical evaluations. The first aims at evaluating the metaphor generation method while the second evaluates the process and the output of the slogan generator method. The last evaluation assesses human-made slogans from similar aspects examined in the second evaluation, to relate the results obtained from both evaluations (on computer-made and human-made slogans) together.

In all evaluations, we run crowd-sourced surveys on Crowdflower[18]. These surveys are targeted to the following English speaking countries: United States, United Kingdom, New Zealand, Ireland, Canada, and Australia.

Table 4 lists the concepts and properties defined by us to evaluate the method. Overall, we had 35 input pairs which will be used as input to both evaluations.

[18] www.crowdflower.com

Concept	Properties	Concept	Properties
book	wise, valuable	chocolate	healthy, sweet
computer	creative, mathematical, powerful	painting	creative, majestic, elegant
car	elegant, exotic, luxurious	university	diverse, valuable
coke	sweet, dark	museum	ancient, scientific
love	wild, beautiful, hungry	professor	old, wise, prestigious, smart
newspaper	commercial, international	paper	white, empty, scientific
politician	powerful, dishonest, persuasive, aggressive		

Table 4: List of evaluated input to the system.

8.1.1 Evaluation of Metaphor Generation

The purpose of this evaluation is to find whether using a metaphor interpretation model to select apt vehicles outperforms selecting vehicles solely based on their strong relatedness with the property (which is considered as the baseline in our evaluation).

In total, the method produces 53 apt vehicles (i.e. vehicles that are considered, by our method, to highlight the input property P in T), of which 31 are general and 22 human, for the inputs defined in Table 4. For each apt vehicle, we select three other vehicles for comparison, as described below. Let *type* denote the type of the apt vehicle, i.e., $type \in \{General, Human\}$.

1. *Apt:* This is the apt vehicle, in the list V_{type} of vehicles considered apt by the metaphor generation method, for which the following three other vehicles are chosen for comparison.

2. *Strongly related:* One vehicle is randomly selected from the vehicle candidates strongly associated with property P (i.e. from top 10% in κ_{type}), but restricted to those that are not considered appropriate by *Meta4meaning* (i.e. not in V_{type}).

3. *Related:* One vehicle is associated with property P but not strongly. It is obtained by picking a random vehicle from the bottom 90% of nouns associated with P in κ_{type}.

4. *Random:* One vehicle is selected randomly among those nouns that are not associated at all with property P in the knowledge base κ.

Examples of generated apt vehicles, for both types of vehicles, along with the other three selected vehicles are given in Appendix in Tables 17 and 18, respectively. Given the 53 apt vehicles, we get 212 metaphors to evaluate overall. For the evaluation, we represent each of them as a nominal metaphor of the form "T is [a/n] v" (e.g., "computer is an artist"). We then asked judges if the metaphor expresses the intended property (that computer is creative). The judges used a 5-point Likert scale where 1 indicates strong disagreement and 5 strong agreement. The order of metaphors was randomized for each judge. 10 judges were required to evaluate every metaphor. For an example of the questionnaire, refer to Figure 8 in the Appendix.

8.1.2 Evaluation of Slogan Generation

We perform this evaluation to identify whether the proposed method is capable of producing expressions suitable for the task, i.e. as advertising slogans. Another goal of the evaluation is to investigate the effects of the evaluation dimensions of the genetic algorithm on the produced slogans. As phonetic aesthetics can be measured computationally, we instead evaluate the effect of prosody features on the catchiness of the expressions. We hypothesise that a balance between the evaluation dimensions of the algorithm is more desirable than maximising some of them. We intend to examine this hypothesis as well.

Below is how we construct the evaluation of expression generation method. For every apt vehicle along with its input, we randomly select two skeletons from the database δ to be filled by the genetic algorithm. We empirically set the following parameters of the genetic algorithm: $\mu = \lambda = 100, G = 25, Prob_c = 0.4, Prob_m = 0.6$.

From the final population produced by the genetic algorithm, we select multiple slogans for evaluation. We select four slogans which maximize each dimension individually. If possible, we randomly select a slogan that has a value on all four dimensions. Additionally, we select two slogans at random where the slogan has positive values on the relatedness and language dimensions, and either of the rhetorical dimensions, at least. Lastly, we select the slogan that has the minimum value on all the dimensions. Some random selections might fail because no slogan in the generated population meets the selection criteria. This selection yields 684 slogans to be evaluated. Finally, to present expressions as in a slogan-like style, we detokenize them using *nltk*[19] [32] and then capitalise the words in them.

[19] www.nltk.org/

We ask 5 judges to evaluate each selected slogan on a 5-point Likert scale based on the following five aspects: (1) the relatedness of the slogan to the title (i.e. input), (2) the correctness of the language, (3) the metaphoricity, (4) the catchiness, attractiveness and memorability, and (5) the overall appropriation of the expression to be used as a slogan. An example of the exact questions asked is given in the Appendix in Figure 9.

8.1.3 Evaluation of Human-Made Slogans

In this evaluation, we run a crowdsourced experiment where we ask online judges to evaluate slogans crafted by professionals. We randomly select 100 slogans from ψ_t. Thereafter, we manually substitute any found product or brand name with "ProductName" to reduce the effect of familiarity of the brand on the evaluation.

We require 10 judges to provide their opinions on the slogans using the same evaluation questions for evaluating computer-made slogans. However, we alter the first question, due to the lack of the input concept and property, as follows. We use the slogan's category provided in the database as the target concept T while removing the property P from the question. Figure 10, in the Appendix, shows an example of the questionnaire.

8.2 Results and Analysis

This section presents the results obtained from the evaluations described above.

8.2.1 Results of Metaphor Generation

We first analyse the results obtained from the evaluation of metaphor generation. We perform three types of statistical analysis. The first illustrates the percentages of judgments on the Likert scale for the different selections. In the second analysis, we provide the mean and standard deviation of the judgements received. Lastly, we perform a permutation test to examine if apt vehicles, as selected by the proposed method, are statistically different from strongly relatedness vehicles.

Figure 3 is a diverging bar chart illustrating the percentages of judgements on the Likert scale for each type of vehicles. We can observe that apt vehicles performed best. Furthermore, quality drops as relatedness strength weakens.

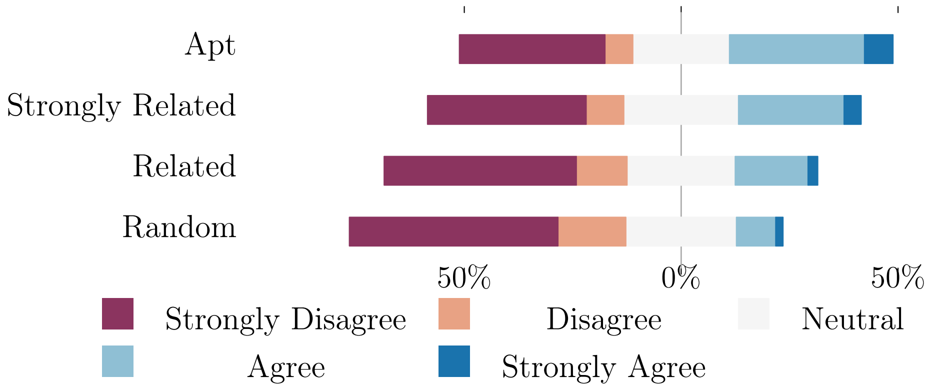


Figure 3: Success of metaphor generation: agreement that the generated metaphor expresses the intended property.

	Apt		Strongly Related		Related		Random	
	μ_x	SD	μ_x	SD	μ_x	SD	μ_x	SD
General	2.51	1.38	2.45	1.30	2.20	1.25	2.01	1.15
Human	2.98	1.33	2.57	1.31	2.22	1.22	2.00	1.08
Total	2.71	1.38	2.50	1.31	2.21	1.23	2.01	1.12

Table 5: The mean and standard deviation of the judgements of metaphors.

Overall, judges agreed or strongly agreed 38% of the time that nominal metaphors constructed with apt vehicles expressed the intended property. On the other hand, metaphors, where the vehicle was strongly associated with the property (but not apt according to the method), were successful in 28% of the cases. The corresponding agreements are even lower for (non-strongly) related vehicles, 19%, and non-related vehicles, 11%.

We next consider the means (μ_x) and standard deviations (SD) of the scores in the Likert scale (Table 5). We also provide these statistics for the two vehicle types evaluated (general and human) vehicles. The number of judgements analysed for each of the four selections (Apt, Strongly Related, Related, Random) is 530, where 310 and 220 of them were general and human vehicles, in the same order.

Based on the statistics, we can observe that apt and strongly related human vehicles, retrieved from V_{Human}, received the highest means, 2.98 and 2.57 respectively, outperforming also apt general vehicles.

The above results show that there is some difference in favour of apt vehicles. We performed a statistical significance test to examine if it is likely that this difference

Concept	Property	Vehicle	Output
computer	creative	artist	Talent, Skill And Support.
			Follow Questions. Start Support.
		poet	Work Unsupervised.
			Younger Than Browser.
car	elegant	dancer	The Cars Of Stage.
painting	creative	literature	You Ca N't Sell The Fine Furniture.
politician	persuasive	orator	Excellent By Party. Speech By Talent.
	dishonest	thief	Free Speech.
	aggressive	predator	Media For A Potential Attack.

Table 6: Examples of generated slogans by the proposed method.

is due to chance. The null hypothesis is that the scores for apt vehicles and strongly related vehicles come from the same distribution, and any difference is due to random effects; the alternative hypothesis is that the mean for apt vehicles is greater than for strongly related vehicles.

We implemented this test as a permutation test, where the two sets of scores were pooled together and then randomly divided into two sets of the original sizes. We ran one hundred million permutations, obtaining an estimate of the distribution between the means under the null hypothesis.

Based on the test, the p-value is 0.0074. The result suggests that apt vehicles perform statistically significantly better than strongly related vehicles.

8.2.2 Results of Computer-Made Slogans

We analyse the results of slogan generation in this section. Table 6 shows some examples of slogans generated by our method. For more examples, refer to Appendix, Tables 19, 21 and 20.

In the following analysis, we consider an individual slogan successful, if the mean score for its overall suitability (the 5th question in the evaluation questionnaire) is above 3. On the average, 35% of generated slogans were considered suitable. The input with the most suitable slogans was, on one hand, *computer–powerful*, with 13 suitable slogans out of 20. On the other hand, the input *newspaper–international* had the least number of good slogans, 1 out of 12. This analysis shows that the

		Relatedness		Language		Metaphoricity		Catchiness		Overall	
Selection	n	μ_x	SD	μ_x	SD	μ_x	SD	μ_x	SD	μ_x	SD
$pos(r, l, m, p)$	262	3.05	0.69	3.15	0.67	2.91	0.60	2.98	0.67	2.92	0.68
$pos(r, l, m)$	93	3.01	0.76	3.06	0.72	2.93	0.61	2.93	0.71	2.87	0.70
$pos(r, l, p)$	111	3.00	0.73	3.17	0.63	2.91	0.63	2.88	0.59	2.86	0.66
$max(r)$	100	<u>3.11</u>	0.70	3.19	0.66	2.90	0.61	2.95	0.68	2.90	0.70
$max(l)$	105	2.89	0.70	<u>3.16</u>	0.70	2.83	0.59	2.91	0.65	2.80	0.68
$max(m)$	88	2.94	0.73	3.01	0.64	<u>2.90</u>	0.62	2.91	0.66	2.83	0.67
$max(p)$	96	2.93	0.76	3.11	0.71	2.91	0.68	<u>2.86</u>	0.67	2.83	0.69
$min(r, l, m, p)$	104	2.77	0.69	2.98	0.65	2.78	0.65	2.82	0.65	2.75	0.70

Table 7: Mean and standard deviation of various judgements of slogans grouped by different selections.

method has successfully generated at least one suitable slogan for each input. Given that the method actually generates an entire population of slogans, more options would be available for an actual user to select from.

Table 7 shows the mean μ_x and standard deviation SD for all slogans evaluated, grouped by the selection methods described in the Evaluation of Slogan Generation section. Letters in the Selection column reflect the four dimensions in the genetic algorithm, i.e. ($\boldsymbol{r}$)elatedness to input, ($\boldsymbol{l}$)anguage, ($\boldsymbol{m}$)etaphoricity, and ($\boldsymbol{p}$)rosody. $pos(*)$ denotes a positive value on all mentioned dimensions only, whereas $min(*)$ and $max(*)$ ensures that they are minimised and maximised, respectively. The number of slogans evaluated for each group is expressed as n.

Observing the overall suitability among all selections, we notice that slogans with balanced dimensions, i.e. $pos(*)$, were appreciated more than slogans with a single dominant, $max(*)$, dimension.

Correctness of the language used in slogans received the highest average rating overall. This is mostly because the language of slogans is checked throughout the entire method (e.g. filling skeletons, mutation, and crossover).

From the examples in Table 6 and opinions on the metaphoricity of generated slogans (Table 7), we can see that the method is capable of generating rhetorical expressions.

Individually maximised dimensions seem to have some correspondence to judgements of their relevant question. For instance, slogans maximising the relatedness dimension, $max(r)$, were judged to be related to the input considerably higher than other selections. However, maximising the relatedness dimension ($max(r)$) seems to increase the scores on all other questions (i.e. language correctness, metaphoricity, catchiness and overall suitability).

Finally, slogans that had the lowest evaluation values on the four dimensions have also received the lowest agreements on all five questions.

We also perform permutation tests on judgements obtained on generated slogans regarding their overall suitability. In this analysis, we divide the data into three sets based on the selection mechanism (i.e. slogans with *balanced* dimensions, slogans with a single *maximised* dimension and slogans with *least* evaluation scores). Using one hundred million permutations, we compare the means under the following alternative hypotheses:

1. $\mu_x(balanced) > \mu_x(maximised)$
2. $\mu_x(balanced) > \mu_x(least)$
3. $\mu_x(maximised) > \mu_x(least)$

Among the tests, only in the second case is the null hypothesis rejected, with a p-value of 0.0286.

These statistics confirm that slogans with balanced values on multiple dimensions (i.e. related to the input, grammatically correct, and have at least one rhetorical device) improve the suitability of slogans, over the case where they are minimised.

Group	n	**Relatedness**	**Language**	**Metaphoricity**	**Catchiness**	**Overall**
Balanced	466	47.64%	52.36%	38.63%	43.78%	39.27%
Maximised	389	45.24%	49.36%	38.56%	39.59%	34.96%
Least	104	27.88%	37.50%	27.88%	35.58%	31.73%

Table 8: The percentage of generated slogans (grouped by the three types of selection) having a given aspect, based on the aggregated judgments received, i.e. having a mean greater than 3.

Table 8 shows the percentage of slogans having a mean judgment greater than 3 on all the considered aspects, divided by the three groups (balanced, maximised and least). In case a slogan has received an average judgment score above 3 on the 5-point Likert scale for a certain aspect, it is considered to be having the aspect. From the table, we notice that the set of balanced slogans contained the most number of slogans having the five examined aspects.

In addition to the above tests, we run four permutation tests. In these tests, we intend to evaluate whether the internal evaluation dimensions (i.e. the four dimensions in the fitness functions) of the algorithm had an effect in improving judgments, on their respective question. In the following tests, we divide judgments into two sets, positive and negative, for each corresponding question to the four dimensions. Positive sets contain individual slogans which received an average judgment on the correspondence question greater than 3, otherwise, they are placed in the negative sets. For example, to evaluate whether slogans that were considered related to target concept and property by online judges were evaluated as related by the method (i.e. on the relatedness dimension to T and P, which contains the two sub-features $f_{rel}(\mathcal{E}, T)$ and $f_{rel}(\mathcal{E}, P)$), we divide slogans based on the judgments received on the first question in the experiment ("The slogan is related to the topic: P T", c.f. Table 9). Thereafter, we run a permutation test under the alternative hypothesis that the mean of the computed relatedness by our method is greater on the positive set than the negative set. We run all the tests using one hundred million permutations. Table 9 shows the results of the tests along with the size of positive and negative sets.

Dimension	**# positive**	**# negative**	**p-value**
Relatedness to T and P	289	395	2.8e-07
Language correctness	325	359	0.8437
Metaphoricity	242	442	0.0456
Prosody	268	416	0.0033

Table 9: Generated slogans were divided into two sets, positive, in case the average judgment on the corresponding question of the dimension was > 3, and negative, otherwise. The number of slogans residing in each set is given in the second and third columns, respectively. The p-values given are the result of the permutation tests under the alternative hypothesis that slogans in the positive set have a mean value on the evaluation dimension greater than the similarly computed mean value of the negative set.

The tests above indicate that our method was capable of estimating three aspects with a statistical significance, which are relatedness to the input, metaphoricity, and prosody. As language correctness received high agreements overall, c.f. Table 7, we believe that the internal evaluation of the language dimension did not have a significant effect due to the multiple language checks employed in the method at different stages.

In the following analysis, we investigate the impacts different skeletons have on the overall suitability of slogans. To do so, we calculate the ratio of suitable slogans (i.e. have an average judgment score greater than 3 on the question concerning the overall suitability) over the total number of generated slogans produced from the same skeleton. Table 10 shows all skeletons, in a simplified form (i.e. showing only tokens and their parts-of-speech in the *Universal POS tags* scheme), used by our method along with the number of total and suitable slogans generated from them. The table points out that distinct skeletons influence the overall suitability differently. As an illustration, the first skeleton in the table has resulted in generating 59% suitable slogans, whilst the last one had only one suitable slogan out of 23. As a result of this analysis, we recommend using the most productive skeletons. Alternatively, if the expression skeletons resource δ contained different skeletons, employing as many skeletons as feasible while generating slogans is advised. This is to acquire diverse slogans and because a given skeleton might perform differently depending on the input.

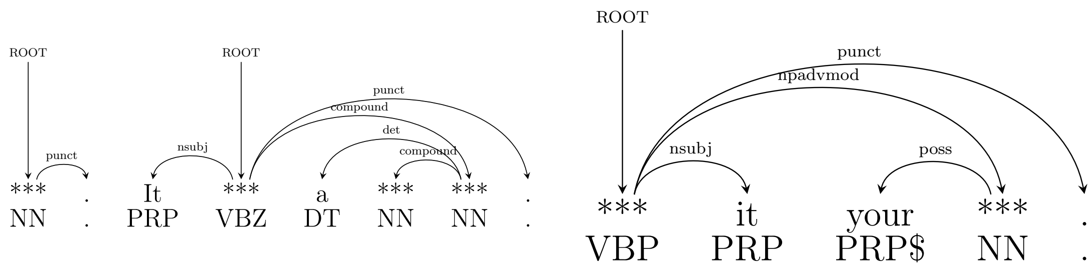


Figure 4: Two skeletons constructed from a metaphorical (left) and non-metaphorical (right) slogans, which are “Success. It’s a mind game.” and “Have it your way.”, respectively.

Next, we look at whether having a skeleton built from an existing metaphoric slogan has an effect on generating metaphorical slogans. To perform the analysis, we manually annotate the 26 unique skeletons based on the original slogan which they were constructed from. Out of the 26 unique skeletons obtained, half of them, 13, were metaphorical (c.f. Table 10). An example of a skeleton constructed from a metaphorical slogan is shown in the left side of Figure 4, the slogan is “Success. It’s a mind game.” by *Tag Heuer*. A skeleton constructed from non-metaphorical slogan, “Have it your way.” by *Burger King*, is given in Figure 4 on the right side. Thereafter, we count how many generated slogans were considered to have metaphoricity in them

Skeleton (simplified)	Metaphorical	# Suitable	Total	Ratio
***_NOUN ,_PUNCT ***_NOUN and_CCONJ ***_NOUN ._PUNCT	Yes	50	85	0.59
***_VERB ***_NOUN ._PUNCT ***_VERB ***_ADV ._PUNCT	No	18	31	0.58
***_ADJ by_ADP ***_NOUN ._PUNCT ***_NOUN by_ADP ***_NOUN ._PUNCT	Yes	24	47	0.51
***_VERB the_DET ***_ADJ ***_NOUN ._PUNCT	Yes	10	23	0.43
***_NOUN for_ADP a_DET ***_ADJ ***_NOUN ._PUNCT	Yes	19	46	0.41
***_VERB ***_NOUN ._PUNCT	Yes	7	18	0.39
***_VERB the_DET ***_NOUN to_PART ***_NOUN ._PUNCT	No	5	13	0.38
***_ADJ than_ADP ***_NOUN ._PUNCT	No	8	22	0.36
***_VERB ***_ADJ ._PUNCT	No	7	20	0.35
The_DET ***_ADJ ***_NOUN is_VERB ***_NOUN ._PUNCT	Yes	8	23	0.35
***_PROPN ***_VERB ***_ADJ ._PUNCT	No	2	6	0.33
***_NOUN ***_NOUN ._PUNCT ***_VERB ***_NOUN ._PUNCT	No	9	27	0.33
The_DET ***_NOUN of_ADP ***_NOUN ._PUNCT	Yes	7	21	0.33
The_DET ***_ADJ ***_NOUN on_ADP ***_NOUN ._PUNCT	Yes	13	40	0.33
***_NOUN never_ADV ***_VERB out_ADP of_ADP ***_NOUN ._PUNCT	Yes	11	38	0.29
***_VERB your_ADJ ***_NOUN do_VERB the_DET ***_NOUN ._PUNCT	Yes	13	48	0.27
You_PRON ca_VERB ***_ADV ***_VERB the_DET ***_ADJ ***_NOUN ._PUNCT	No	8	31	0.26
***_VERB ***_NOUN ***_NOUN ._PUNCT	No	6	24	0.25
***_PROPN ***_ADV ._PUNCT	No	4	18	0.22
***_VERB ***_NOUN the_DET ***_NOUN over_ADV ._PUNCT	No	3	16	0.19
***_NOUN ***_VERB and_CCONJ ***_VERB and_CCONJ ***_VERB ._PUNCT	No	1	6	0.17
It_PRON ***_VERB ***_NOUN ._PUNCT	No	3	19	0.16
Between_ADP ***_NOUN and_CCONJ ***_NOUN ***_VERB ***_NOUN ._PUNCT	Yes	2	13	0.15
***_VERB ***_NOUN ._PUNCT	Yes	2	14	0.14
I_PRON ***_VERB ***_VERB it_PRON ._PUNCT	No	1	12	0.08
***_NOUN ._PUNCT It_PRON ***_VERB a_DET ***_NOUN ***_NOUN ._PUNCT	Yes	1	23	0.04

Table 10: The number of generated slogans (suitable and total) for every skeleton employed in our generator, sorted descendingly based on the ratio of suitable slogans to total slogans. The second column indicates whether the skeleton was produced from a metaphorical slogan.

in regard to the skeleton type, metaphorical or not, given in Table 11. In total, 35%, 242, of all generated slogans were considered to have metaphoricity in them. Out of the generated slogans, 38% of them were based on a metaphorical skeleton and considered to contain metaphoricity, whereas 31% of them were based on a non-metaphorical skeleton and considered to contain metaphoricity. These results indicate that generating slogans using skeletons of slogans conveying a metaphor initially has a higher potential to produce metaphorical slogans as well. Despite that, our proposed method appears to be capable of generating metaphorical slogans from non-metaphorical skeletons.

8.2.3 Results of Human-Made Slogans

In this section, we analyse the results obtained from evaluating existing slogans. Overall, we received 1,000 judgments on the evaluated 100 slogans, 10 judgments per slogan. Throughout the coming analyses, we aggregate judgments of a given slo-

Skeleton Type	**Metaphoricity**	**No-metaphoriciy**	**Total**
Metaphorical	166 (38%)	273 (62%)	439 (100%)
Non-metaphorical	76 (31%)	169 (69%)	245 (100%)

Table 11: The number of generated slogans that were deemed to have metaphoricity, divided based on whether their skeleton was formed from a metaphorical slogan.

gan by their mean. We perform these analyses to relate the results of human-made slogans to the ones obtained from evaluating computer-made slogans and to reveal interesting insights (e.g. correlations between the aspects considered) to be considered in the future research. It is worth noting that a direct comparison between the results of computer-made slogans and human-made ones is not feasible. This is due to various reasons such as the difference between the two evaluations (e.g. missing the adjectival properties from evaluated human-made slogans, nonequivalent number of judges and nonidentical judges) and the conditions enforced during producing the slogans (e.g. computer-made slogans were restricted to a skeleton per generation), in addition to the fact that generated slogans are intended to be slogan candidates (i.e. not finalised). Nevertheless, in most of the following analyses, we juxtapose the results of analysing existing slogans with computer-made slogans for finding interesting insights.

Table 12 illustrates the mean and standard deviation of the received judgments. We can observe that most of the existing slogans contained the examined aspects such as relatedness to the product and overall suitability, i.e. have received a mean > 3. Existing slogans received an average agreement score of 3.63 for being good slogans.

		Relatedness		**Language**		**Metaphoricity**		**Catchiness**		**Overall**	
Made by	n	μ_x	SD	μ_x	SD	μ_x	SD	μ_x	SD	μ_x	SD
Humans	100	3.77	0.49	3.99	0.34	3.38	0.37	3.60	0.38	3.63	0.40
Computer	262	3.05	0.69	3.15	0.67	2.91	0.60	2.98	0.67	2.92	0.68

Table 12: Mean and standard deviation of judgments of existing slogans (in the first row). For comparison, the statistics of the same analysis but on the computer-made slogans having a balanced value on all dimensions (c.f. Table 7) are given in the second row.

The language correctness of the sampled slogans received the highest agreement and had the least standard deviation. This suggests that slogans made by professionals do not have usually have grammatically wrong expressions. Furthermore, it indicates that language correctness is essential for constructing slogans.

Many slogans were considered related to the product by the judges, based on the mean 3.77. The standard deviation of the relatedness score is the highest, which suggests that judges had different opinions regarding the strength of the relatedness.

On average, existing slogans seem catchy and attractive, based on the received judgments. The mean score received for catchiness was 3.6. The high number of catchy slogans observed imply that catchiness is a role-playing element in creating suitable slogans.

Finally, existing slogans were considered to be metaphoric, but with an average score closer to neutral than the remaining evaluated aspects, 3.38. This score is justifiable by the fact that metaphoricity is not a strong requirement for producing successful slogans; however, the score hints that more than half of the slogans were found to be metaphoric.

Looking at the difference of means (and standard deviations) of judgments on the different aspects between human-made and computer-made slogans in Table 12, we see that human-made slogans had higher scores and agreements. Based on the differences of means, we see that computer-made slogans have received the closest judgments to human-made on the metaphoricity aspect (with 0.47 mean difference) and the furthest on the language aspect (with 0.84 mean difference).

Made by	n	**Relatedness**	**Language**	**Metaphoricity**	**Catchiness**	**Overall**
Humans	100	94%	98%	84%	89%	92%
Computers	466	48%	52%	39%	44%	39%

Table 13: The percentage of slogans having a given aspect, based on the aggregated judgments received, i.e. having a mean greater than 3. The results of the (balanced) computer-made slogans are added for reference, see Table 8 for the results of other types of computer-made slogans.

To further analyse the results, we look at the percentage of slogans that were considered to have a certain aspect, see Table 13. Similar to the analysis conducted for computer-made slogans (c.f. Table 8 for the results), if the average judgment score for a slogan was above 3 on the 5-point Likert scale for a given aspect, the slogan is regarded as having the aspect.

In terms of relatedness to the product, 94% of slogans were considered to cover this aspect, which is a huge portion. Random examples that have this property are "Go new places." by *Atlas Van Lines* –a moving company– and "Everyone's cup of tea." by *Tetley*. Examples of slogans that were not believed to be related to the product are *Head & Shoulders*' "Stay faithful." and *The Sum*'s, a graphic design studio, "Do Sumthing good.". These two negative examples do not easily highlight the advertised product or service, whereas the positive examples clearly indicate them.

Looking at the statistics, we observe that almost all slogans, 98% were considered to be grammatically correct. The two slogans that were considered to have mistakes are "Highly REDommended." and "Do Sumthing good." by *RED* –a driving school– and *The Sum* –a graphic design studio–, respectively. Despite that these slogans were considered to have incorrect language, they are clearly puns. As the brand name was not provided in the survey (manually replaced with the placeholder *ProductName*), judges were unable to spot these puns. This confirms that languages correctness is indeed an essential element in constructing slogans.

84% of slogans were credited as metaphoric. Out of all the other aspects, metaphoricity is the least frequent one; although, common. "That frosty mug sensation." by *A&W; Root Beer* and "Expand your world." by *Westbank Library* are randomly selected examples of slogans that were considered metaphoric. On the other hand, "Wine. What are you saving it for?" by *Wine Market Council of California* and "ProductName. Serve Chilled." by *Baileys Irish Cream Liqueur* are examples of non-metaphoric slogans, from the perspective of the judges.

Cathiness existed in 89 slogans out of 100. "ProductName. Awaken the spirit" by *Kahlua* –a brand producing Mexican coffee liqueur– and "Taxi! Taxi!" by *Santa Monica Taxi* are examples of catchy slogans. Some examples of slogans which judges did not find to be catchy are: (1) "The longer lasting pleasure." by *Haagen Dazs* (an ice-cream shop) and (2) "Every body could use some good ProductName." by *Karma Wellness Studio*.

Overall, we notice that 92% of the slogans were considered suitable. The slogans by *Seat-Ibiza Automotive* and *Raleigh Home Painters & Tri Son Services*, "ProductName. Different rituals, same spirit." and "Adding color to your life!", in the same order, are examples of suitable slogans. On the contrarily, *Haagen Dazs*'s and *The Sum*'s slogans were pointed out as non-suitable.

The above percentages indicate that the four aspects (relateness to the product, language correctness, metaphoricity and catchiness) to an extent, which impacts their suitability for the task (i.e. as advertising slogans). In terms of the negative cases, the evaluation setup might have affected the judgments as the brand/product's name was not provided and, if it appeared in the slogan, it was substituted with a placeholder. Nonetheless, four non-suitable slogans do not fall under either of the situations, such as "Unequalled luxury." by *Asanti Wheel* and "Whatever your financial requirement we have a solutions for you." by *OPM Leasing*.

The next analysis we perform is the inter-rater reliability using Krippendorff's alpha. For each question in the experiment, we perform the analysis by calculating the reliability of all individual judges on all slogans, given in Table 14. The results of the inter-rater reliability suggest that judges disagreed frequently while judging slogans. These results exhibit the subjectivity of opinions when perceiving human-made creative expressions.

Relatedness	**Language**	**Metaphoricity**	**Catchiness**	**Overall**
0.1179	0.0463	0.0144	0.0309	0.0565

Table 14: The inter-rater reliability among all judges for each question.

To obtain further understanding of how the four dimensions are associated with each other and their influence on the overall suitability of slogans, we calculate the correlations of judgments between the five questions. We utilise Pearson correlation coefficient to measure linear correlations. Figure 5 is a heatmap of these correlations.

Observing the overall suitability of slogans, we notice that the remaining four questions correlated positively with it, which indicates that their existence improves slogans. Catchiness and relatedness to the product had the highest correlation with suitability, making them also essential for producing good slogans.

Generally, all questions had a positive correlation with each other. Additionally, all of the correlations were strong, i.e > 0.5, except two correlations which are metaphoricity and language correctness, and metaphoricity and relatedness. The low correlation to relatedness could be because metaphoricity occurs when another concept, the vehicle, is expressed in the slogan; hence, a decreased relatedness to the input concept. Nevertheless, metaphoricity had a strong positive correlation with catchy and good slogans.

The different levels of correlation between the questions suggest that a sufficient balance of these aspects in slogans is needed to produce successful slogans.

Next, we plot the distributions of the aggregated results obtained on the five ques-

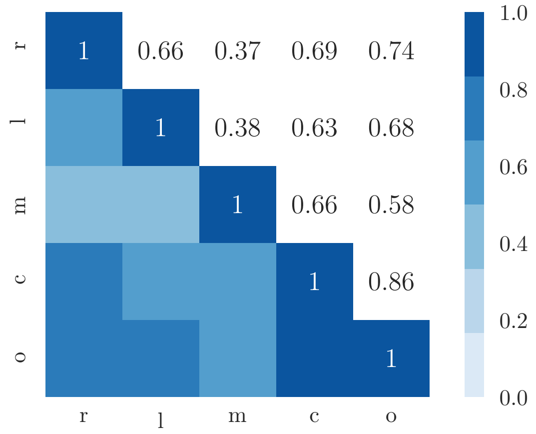
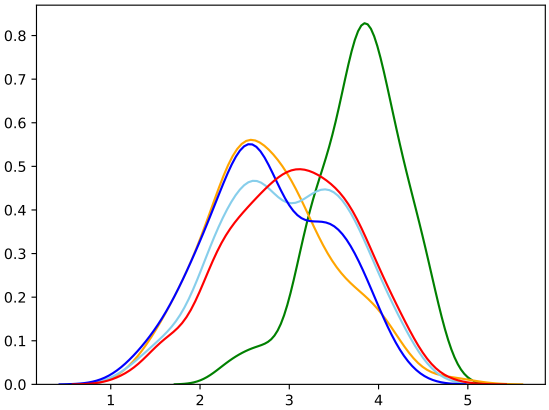
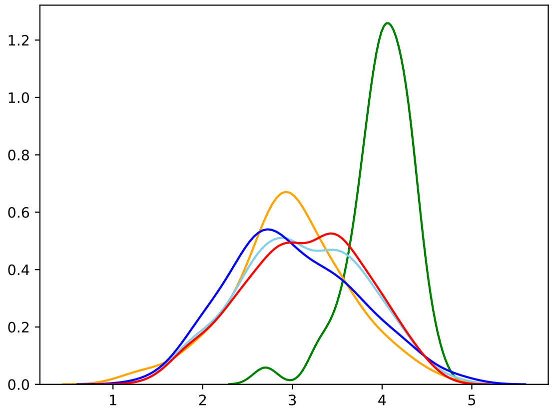


Figure 5: Pearson correlation coefficient of judgments between the five questions: (r)elatedness, (l)anguage, (m)etaphoricity, (c)atchyness and (o)verall suitability.

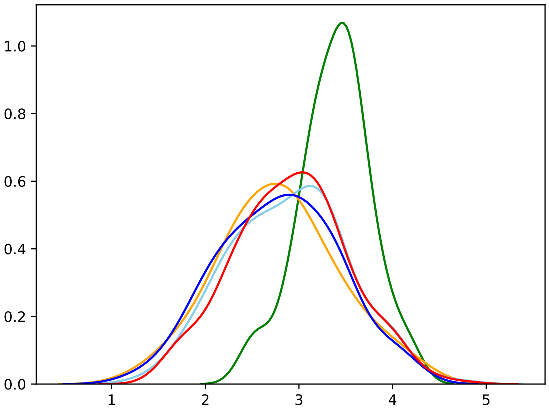
tions from both slogan surveys, i.e. results of computer-made and human-made slogans. We perform this analysis to visualise the performance (based on a summary of the distributions) of the proposed method for generating slogans in relation to human-made slogans. In this analysis, we consider the following types of computer-made slogans: balanced, maximised, maximised without any balanced and least. The distributions are shown in Figure 6, the x-axis represents the 5-point Likert scale whereas y-axis indicates the normalised number judgments. From the plots, we notice that the distribution of balanced slogans seems to contain more positive (i.e. residing within $(3, 5]$ on the x-axis) and less negative slogans than the other distributions of computer-made slogans. Also, we recognise that human-made slogans were deemed to possess the examined aspects, as most of the distribution lies on the right side of the x-axis (i.e. > 3). However, the degree of skewness of human-made slogans on the five aspects varied (e.g. the distributions on the language correctness and metaphoricity aspects). Focusing on the balanced computer-made distributions, we see that they are relatively symmetrical. Such symmetricity indicates that most slogans received neutral judgments, while having some slogans with agreements and disagreements on all aspects. The distributions of judgments obtained on the overall suitability of the least and both maximised types of computer-made slogans are concentrated between $[2 - 3)$ on the x-axis, suggesting that a significant number of their slogans were found to be non-apt for the task. Observing the two types of distributions, we discern that the slogan generator needs further development and research to reach a state similar to professionals advertisers in producing slogans. Despite that, based on the results provided in the section "Results of Computer-Made Slogans", the generator has successfully generated suitable slogans for every



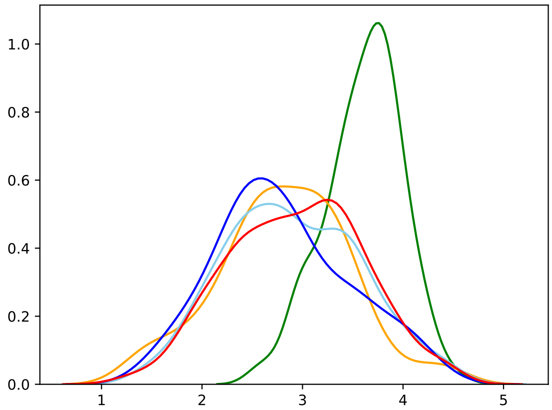
(a) Relatedness



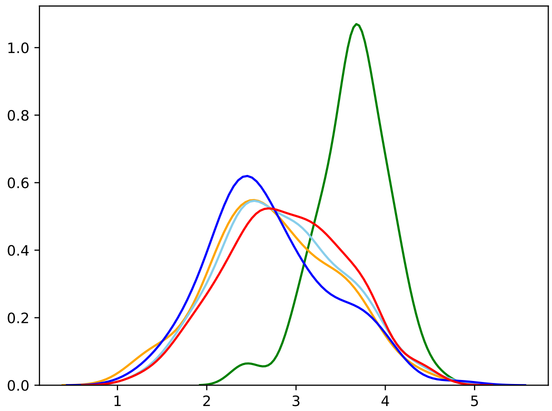
(b) Language



(c) Metaphoricity



(d) Catchiness



(e) Overall

Figure 6: Distributions of judgments received on existing slogans (in green) and generated slogans (*balanced* in red, *maximised* in light blue, *maximised* excluding any *balanced* slogans in blue, and *least* in orange).

input concept.

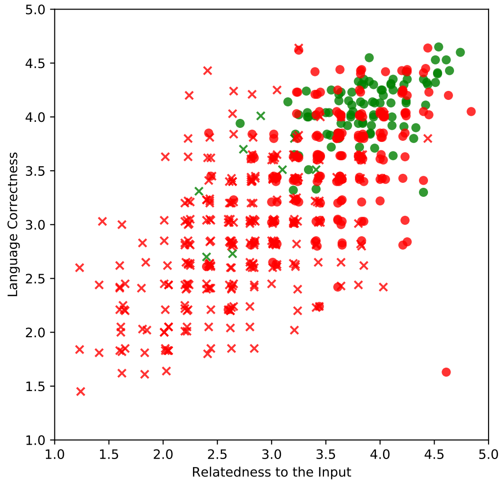
In our next analysis, we show scatter plots in Figure 7 of man-made and balanced computer-made slogans for every combination of two aspects, excluding overall suitability. As to our prior results on correlations between the considered aspects, the plots also exhibit the correlations. Generally, the plots reveal that the overall suitability of slogans increases at times where the evaluated aspects exist in them. Notwithstanding, some pairs of aspects do not clearly reflect a positive effect on overall suitability, for computer-made slogans. For instance, some computer-made slogans in the two pairs language-catchiness and metaphoricity-catchiness were considered to be non-suitable despite having high scores on both dimensions.

9 Discussion

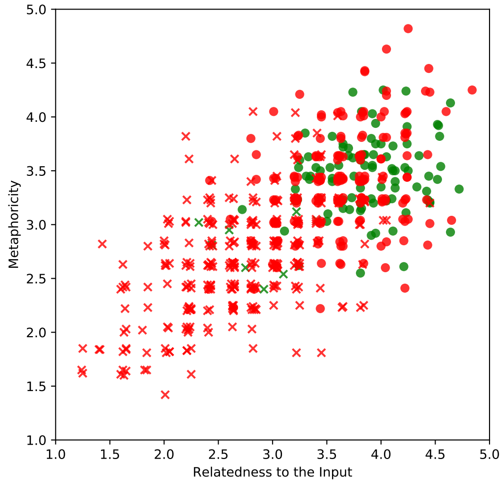
In this section, we discuss the work conducted in this thesis, along with the results. We, also, relate them to previous research when feasible.

In our manual analysis (section 5), we have annotated 100 randomly selected human-made slogans where we have classified slogans based on their usage of rhetorical devices, following the taxonomy proposed by Mcquarrie and Mick [36]. The results of the analysis verify that slogans frequently employ various figurative devices, as also concluded in previous studies [40, 37]. By following the taxonomy provided by Mcquarrie and Mick in our manual analysis, we organised slogans structurally and hierarchically based on their rhetorical devices. However, a limitation of following the taxonomy is that it does not cover all types of rhetorical devices (e.g. aphorism). Such a limitation was not crucial for our research as the taxonomy sufficiently included many common rhetorical devices.

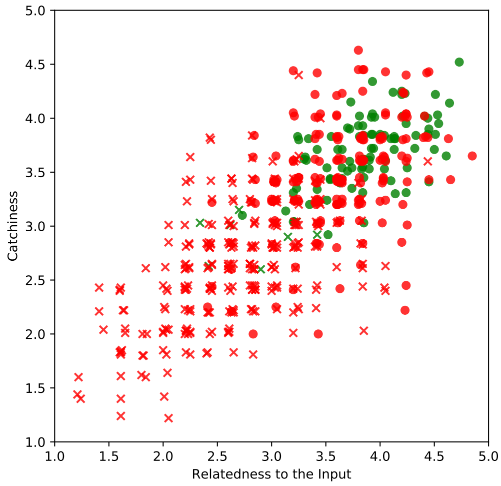
From our manual analysis, we noticed that the most observed rhetorical operation, alone and in combination with other operations, is repetition (e.g. sounds and words). This could be due to the simplicity of creating and noticing such repetitions along with their benefit of increasing the recall of the brand. Destabilization is the second most seen rhetorical operation in slogans, alone and with other operations. Despite the rhetorical operation being considered as complex (as it greatly deviates from the audience's expectation which makes constructing and comprehending such operations require additional effort), we believe it had such high occurrence because it demonstrates higher creativity, in comparison to the repetition rhetorical operation for instance. Having a slogan with high creativity makes it stand out which



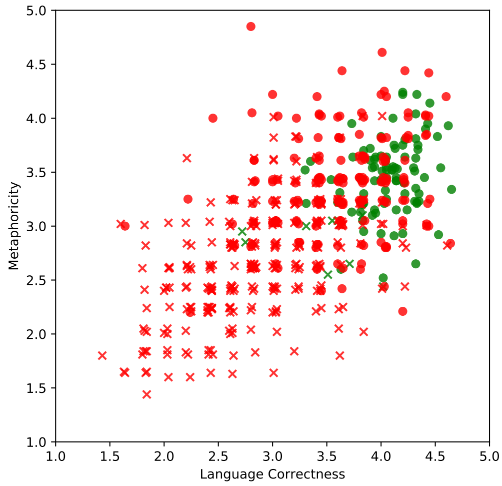
(a) Relatedness vs Language



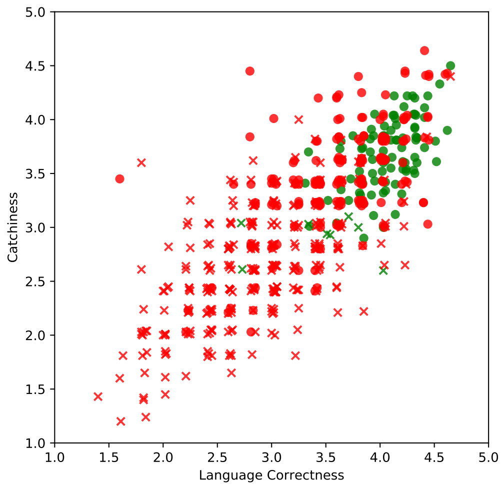
(b) Relatedness vs Metaphoricity



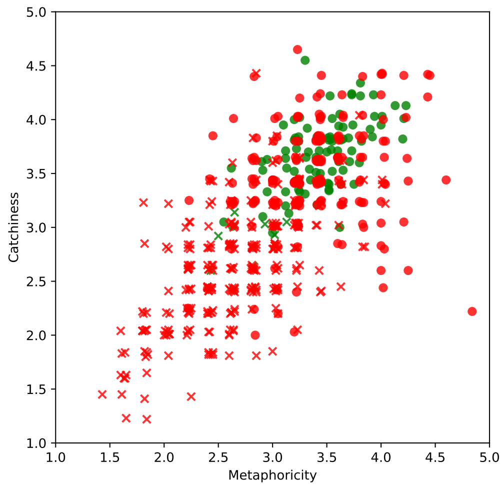
(c) Relatedness vs Catchiness



(d) Language vs Metaphoricity



(e) Language vs Catchiness



(f) Metaphoricity vs Catchiness

Figure 7: Scatter plots illustrating suitable and non-suitable human-made (in green) and balanced computer-made (in red) slogans, for a pair of aspects. Slogans that have received an average overall suitability greater than 3 are indicated by a filled circle, otherwise by an X.

results in high recognisability. 32% of slogans had a combination of rhetorical operations (where a given operation might contain multiple rhetorical devices). Similarly, this could be employed in slogans by advertising professionals to achieve higher creativity. This result motivates considering and balancing as many rhetorical operations/devices as possible while constructing slogans, manually or automatically.

Performing the manual analysis by one annotator on 100 random slogans is not adequate for generalising our findings. This is because the sample might contain more frequent usages of certain rhetorical devices by chance. Also, the task of annotating usages of rhetorical devices in slogans is subjective (e.g. should the metaphoricity be judged based on the slogan itself or while considering the product/brand as well). Therefore, similar additional annotations by multiple experts are needed to confirm a generalisation of our results. Lastly, our manual analysis did not take into account the popularity of the slogans, which prevents us from knowing how effective it is to utilise a given set of rhetorical devices. Nevertheless, the popularity of a slogan could be estimated by looking up the number of times it occurred on the Internet by querying *Google Search API*.

The goal of our semi-automatic analysis, conducted in section 6, is to evaluate the performance of different NLP tools and services in processing/predicting multiple aspects of slogans. From the results, it is evident that the considered tools did not achieve high accuracy, mainly in predicting the categories of slogans and detecting any named entities used in them. As slogans are usually informal and short, most of the current NLP tools did not parse them correctly. However, *Spacy* had the highest parsing accuracy, 88% of the slogans, which is sufficient for our case. These results suggest that informal sentences are usually overlooked by existing tools and more research is needed to enhance them. High accuracy, $\geq 93\%$, of correct predictions was achieved by the two examined sentiment analysis tools during our manual analysis. However, our test data (slogans) is biased as it only contains expressions with positive sentiment. To assess these tools properly, negative examples should have existed in our data.

There are multiple linguistic resources and tools utilised in our method, e.g. the text corpus and the language and semantic models. Deficiencies in any of them may influence the output and results of the method. For instance, in some cases, the natural language processing tool might fail in parsing sentences correctly. Such failures result in building incorrect grammatical relations and skeletons which, then, affect the grammatical correctness of generated expressions. Using misleading knowledge from the resources causes generating grammatically incorrect slogans.

Some of the functions defined by us for measuring linguistic characteristics (in section 7.2) could be implemented differently to handle the same criteria. For example, $f_{grammar}$ could use a rule-based approach to check the grammar of expressions, instead of the grammatical relations repository γ. Accordingly, the resources, evaluation functions and the proposed method could be modified to resist similar failures. An example of such modification is introducing a layer in the slogan generation method for validating and correcting the grammar of generated slogans (e.g. a rule-based grammatical correction method for correcting obvious grammatical mistakes such as "I're").

During our computational analysis of slogans 7.3, we evaluated the functions defined by us on existing slogans in the corpus ψ_t. Albeit evaluating most of the functions, we were unable to evaluate the metaphoricity function (i.e. $f_{metaphoricity\_1}$ and $f_{metaphoricity\_2}$). This is because an adjectival property P of the product (intended to be highlighted in the slogan) is required as a parameter to the functions but is missing from the corpus. Despite that, the results of the permutation tests conducted in section 8.2.2 (shown in Table 9) infer that these functions, combined, have estimated the metaphoricity of slogans with a statistical significance. Regardless of this statistical significance, there is no guarantee that the metaphorical slogans produced by the method convey the intended metaphor. A way of evaluating this aspect would require us to ask the judges to provide their reasons for finding the slogans metaphorical. The results of such questionnaire would then be qualitatively analysed to assess this question. A mean for increasing the possibility of conveying the intended property is to also consider it in the metaphoricity dimension.

The function for measuring the semantic cohesion between words in an expression, $f_{cohesion}$, found out that only a small amount of slogans (15%) had semantic cohesion. We believe this is because of the strict condition that all words should be related to each other, and because of missing individual pairs from the semantic model. Having a more lenient condition (e.g. some of the words should be related to at least one word) would be more realistic to enforce.

Based on the functions ($f_{assonance}$, $f_{consonance}$, $f_{alliteration}$ and f_{rhyme}), 9%, 14%, 11% and 13% of the automatically analysed slogans contained assonance, consonance, alliteration and rhyme, respectively. On the contrary, our manual analysis shows that these features existed in the following percentages of slogans: 16%, 26%, 14% and 9%, in the same order. From both results, it appears that consonance is the most used prosody feature. The increased detection of rhymes by f_{rhyme} could be caused by matching on one phoneme, instead of syllables. To further evaluate the proposed prosody functions, we could evaluate them on the manually annotated slogans and compare their results to our annotations.

In this thesis, we have proposed a method for generating metaphors, a figurative device that is very often found in successful slogans. The metaphor generation method employs a metaphor interpretation model –*Meta4meaning*– to measure the aptness of vehicle candidates. We have compared the method against metaphors generated based on strong relatedness to input property. The results of the evaluation indicate that using a metaphor interpretation model produces better metaphors. Nevertheless, as the metaphor generation method relies mainly on *Meta4meaning*, a failure of interpreting a metaphor by the model for any of its limitations, c.f. Xiao et al. [53], might treat apt vehicles as non-apt.

The proposed method for slogan generation needs multiple parameters to be tuned to achieve better results (e.g. genetic algorithm parameters). Altering these parameters could improve the results. Reducing the number of related words to consider shrinks the interesting search space which would result in generating few, yet productive, slogans. On the other hand, increases in the space might result in including non-related words as interesting word candidates, i.e. generate more poor slogans. As relatedness to the product (and/or property) seems to be very important in producing (by our method and humans) suitable slogans, based on the sections 8.2.2and 8.2.3, ensuring that the interesting search space contains enough related words of a good quality is essential to producing diverse and suitable slogans. Another example for adjusting the parameters is increasing the population size μ and the number of generations G, which would result in searching a wider range of the search space but would, also, increase the generation time.

From our analysis of the results of computer-made slogans, it appears that the language dimension did not have a significant effect on the correctness of the language, refer to Table 9. We believe that this is because grammatical relations are checked throughout the process (e.g. during mutation and crossover). Hence, dropping the language dimension might increase the performance of the method as it would concentrate on the essential dimensions. Another feature that might enhance the performance of the method when removed is the coherence function. As we have employed a strict coherence function (i.e. all words must be related to each other) in the filtering phase, suitable candidates are possibly filtered out due to a failure in satisfying this condition. It would be beneficial, in future research, to evaluate different variations of the proposed method where a minimum base system (e.g. having the relatedness dimension only without any filtering) is compared to the same base with added single and multiple combinations of components (e.g. interesting space, metaphoricity and catchiness). Such evaluations would examine the implications of considering each component of the overall quality of slogans.

In section 8.1 we have described our evaluation setup for evaluating the proposed method and existing slogans. In our evaluations, we have asked ordinary people on the Internet, using a crowdsourcing platform, to judge metaphors and slogans. We chose to obtain opinions of ordinary people, in contrast to experts, because they are easier to reach; although, hiring experts would yield a higher quality of results. A hard-to-mitigate risk of running crowdsourced subjective tasks that do not have a fixed solution is scammers which abuse the system by randomly solving the task without paying attention to maximise their income. Increasing the number of judges would assist in finding a common opinion for a given question and conduct statistical analysis with a higher confidence; however, doing so would also increase the financial costs of the experiments. Furthermore, for such subjective tasks, the sufficient and acceptable number of judges is difficult to identify. As a result, we estimated it depending on the amount of data and costs.

In addition to the above challenges, a difficulty we faced during our evaluation setup is defining questions that reflect our evaluation goals while reducing plausible confusions by online judges. This is because we had to assume that online judges have no knowledge of any specific linguistic concepts such as metaphors, semantic relatedness and prosody. Regardless of our best efforts in crafting the questions, it is infeasible for us to verify that the judges have actually understood the task fully and answered accordingly. The standard deviation SD of judgments received on computer-made and human-made slogans, shown in Table 12, we notice that the relatedness question has the highest standard deviation in both cases, followed by the overall suitability. Such deviations show that judges did not have a consistent opinion, which could be due to different understandings of the question or representations of agreements. Converting the judgments received into binaries helped tackle such diversity.

The evaluation of human-made slogans indicates that the evaluated aspects (i.e. relatedness to the product, correctness of the language, metaphoricity and catchiness) contribute to the overall suitability of the slogan. Juxtaposing the results of human-made with computer-made (balanced, especially) slogans, we see that some generated slogans received scores similar to existing slogans (c.f. distribution plots in Figure 6). Additionally, the percentages of computer-made slogans having a certain aspect relatively match the results of human-made slogans, given in Table 13. For instance, the language correctness and relatedness aspects were deemed to exist the most in both types of slogans, respectively, while metaphoricity was considered to exist the least in them. The same can be observed in the mean of judgments provided in Table 12. Interestingly, from the results, it appears that multiple computer-made slogans were considered catchy but not suitable, which is not the case for human-made slogans.

Nevertheless, there were many low-scored (in terms of the overall suitability) generated slogans, in comparison to man-made slogans, which articulates that further research is required to achieve generating high-quality slogans. Despite that, the intended use-case of the proposed method is to generate inspirational slogans to be used during brainstorming sessions, where professionals would interactively adjust the method (e.g. input, skeletons and weights of the fitness functions) and selectively choose good slogans.

A different and meaningful evaluation for measuring the effectiveness of the proposed method in generating slogan candidates is requesting online workers (easier to recruit) to produce multiple slogans having the same skeletons, target concepts and adjectival properties (Tables 4 and 10) as used by the method. Furthermore, workers could optionally state why and how they came up with the slogans. Conducting such an evaluation and qualitatively analysing its results would help us find common aspects and links that people tend to consider when suggesting slogans for certain concepts. This knowledge facilitates modelling the process of constructing suitable slogans computationally, especially if the workers were advertising professionals. This evaluation could be followed by another one where different ordinary people (acting as the audience) assess the quality of slogans produced by the human workers. This corpus (i.e. concepts and adjectival properties along with their human-made slogans and their evaluation scores) would then act as a baseline for the creative task of suggesting inspirational slogans, assuming that the slogans suggested by the workers are representative examples of good slogans such as those that could be proposed in an advertising brainstorming session. The baseline could be, then, compared directly to the results obtained in our evaluation of computer-made slogans. This would, also, make it possible to evaluate the slogans output by different slogan generators. This evaluation is left for future work.

The purpose of the proposed method is to act as an auxiliary tool for professionals when constructing slogans. Our evaluations did not cover evaluating the method in this use-case. To evaluate how the method would aid in inspiring professionals while generating slogans, we would need to provide the method as a user-friendly service/tool for them to use during a brainstorming session. Once the session is over, professionals could provide their assessments of the value of the service. Additionally, the service could continuously monitor any adjustments of parameters and slogans selected/saved by the professionals, to estimate the effects of modifying the parameters on generating good slogan candidates. Such an evaluation is planned for future work, as it is beyond the scope of this thesis.

In our evaluation, we did not evaluate the creativity of the method nor the creativity of its output, as we assumed that the task of producing suitable slogans is creative; hence, a method capable of doing that should be perceived as creative. To evaluate the creativity of the method, the procedure proposed by Jordanous [27] could be followed, as shown in the work by [1] for evaluating another creative task. In short, the procedure expects us to define what is creativity in general and in our particular case, and based on our definitions, derive the evaluations.

We hypothesise that a possible generalisation from our research is that considering multiple dimensions and balancing them would be beneficial for producing other creative artefacts, linguistic (e.g. poems, songs and news titles) or non-linguistic (e.g. paintings and music). If we were to bring this generalisation into the context of a particular problem such as that of generating poems, some of the features considered in our method (e.g. prosody, relatedness to the topic and metaphoricity) along with additional ones (e.g. meter and poem-wide semantic coherence) could be employed and balanced during the generation process. As a poem consists of multiple smaller units of a length of a slogan, the complexity increases. Hence, it might be necessary to balance the features on multiple sub-structural levels such as verse-, stanza- and poem-level to produce diverse, yet connected verses.

10 Conclusions

In this thesis, we have described a method for generating slogans computationally. We have evaluated the proposed method by running crowdsourced questionnaires. The evaluation showed that the proposed method generates suitable slogans, given a target concept (e.g. car) and an adjectival property to express in the slogan (e.g.

elegant). Not only that but, as part of the method, metaphors were generated and used in producing metaphorical slogans. For instance, for the earlier given input, the method generated "dancer" as a suitable vehicle to express the property *elegant* in *car*, which was then used to generate the slogan "The Cars Of Stage.". Parts of the conducted research in this thesis have resulted in a publication [3] and were employed in another creative task automatic generation of satire [1].

Prior to constructing our method for generating slogans, we have studied the related literature and conducted manual and semi-automatic analyses to understand previous work on tackling the issue and obtain a detailed understanding of slogans and their linguistic characteristics. We used this knowledge in defining computational functions for estimating potentially crucial linguistic characteristics of slogans and proposing the method for generating metaphors and slogans.

Our method for generating slogans is based on genetic algorithms using multi-objective selection. The method has successfully created slogans that were considered suitable, related, grammatically correct, metaphorical and catchy, based on crowdsourced opinions. Furthermore, the method was capable of generating metaphorical expressions beyond restrictive metaphorical templates (e.g. nominal metaphor "T is a/n v" and similes "T as P as a/n v"), e.g. "The Cars Of Stage.".

We have conducted multiple experiments and analyses to evaluate different aspects of existing slogans and the proposed method. For example, we have evaluated the defined functions for measuring linguistic characteristics on a big corpus of existing slogans to underline their presence in slogans. Furthermore, in the experiments, we have assessed (by asking online judges) the outputs of the method for generating metaphors and slogans, along with existing slogans. The results of the experiments were thoroughly analysed to evaluate the method.

A summary of the results of the research conducted in this thesis is the following. Rhetorical devices are employed in many slogans, based on our manual and automatic analysis. We have analysed the results of the three crowdsourced evaluations (on metaphor, and computer-made and human-made slogans) from various viewpoints. From the evaluations on the generated metaphors and slogans (from the metaphoricity aspect), it is evident that the method was able to create metaphors and metaphorical slogans automatically. Concentrating on the evaluation of computer-made slogans, we observe that generated slogans incorporated relatedness to the input, correct language, metaphoricity and catchiness. Most importantly, we notice that slogans with balanced features (i.e. having a positive value on multiple fitness functions) outperformed slogans maximising a singular dimension. Further insights were disclosed by our evaluations such as the effect of employing different skeletons on producing suitable slogans and using metaphorical skeletons on generating metaphorical slogans. Comparing the results of computer-made and human-made slogans, we see that the method has produced slogans with similar scores to existing ones.

All the research questions in this thesis were addressed throughout our work, e.g. analyses (manual, semi-automatic and automatic) and empirical evaluations. Our manual analysis contributes mainly in answering the research questions RQ2.1 and RQ2.2, and partially in answering RQ1.1-3. Also, it answers some elements of the first research question, about commonly used linguistic characteristics in slogans, partially. It does so by revealing the consonance and assonance features that were found in the slogans, which addresses RQ1.1 (the phonological characteristics of slogans). Syntax (RQ1.2) and semantics (RQ1.3) were also covered by our manual analysis, in the case of figurative usages, e.g. parison and paradox, respectively.

In our semi-automatic analysis, we have covered the remaining linguistic characteristics from the first research question, such as sentiment (RQ1.4), semantics (RQ1.3) by predicting the category, and references of the product/brand name in the slogan. It also provides answers to the question of which computational tools and resources augment the generation of slogans, i.e. RQ3.3. The research questions concerning computationally modelling rhetorical devices and other linguistic characteristics (RQ2.3 and RQ3.1-3) of slogans is tackled in the computational resources used throughout the thesis and the functions for measuring such characteristics.

Our metaphor generation method along with its employment in the slogan generator answers the RQ2.3. The method for generating slogans deals with the research question with respect to how can we generate slogans similar to slogans produced by humans (specifically on syntax, aesthetics and semantics), which are RQ3.1-2.

The fourth research question (RQ4) is fulfiled in the crowdsourced evaluation of computer-made slogans. Our empirical evaluation covered the remaining research questions, which are RQ4, RQ5.1 and RQ5.2. Notwithstanding, future research is needed to improve the method and test it in production.

A possible future direction for the metaphor generation method is introducing additional factors to rate the aptness of generated vehicles, e.g. shared categories between tenor and vehicle, as they could improve the aptness of the generated vehicles. For instance, vehicles that do not share a category with the tenor could be omitted. However, additional factors would typically require more knowledge bases to be used. A possible future direction for metaphor generation is to combine an interpretation model with additional measurements to reach aptness scores matching how humans perceive metaphors.

Studying the effects of adjusting the parameters of the method on the results is left for future work. These parameters could be altered dynamically based on the interactions between the user and the system, which would motivate collaborations between humans and computers in solving creative tasks.

Finally, the proposed method for generating slogans could be adapted to/for other creative writing domains that have contextual texts, as slogans tend to have no textual context other than the brand/product. For example, it could be used in generating attractive news titles (given a textual news article) by using skeletons of news titles and introducing additional meta-information to handle cases of covering multiple concepts and details (e.g. comparing two parties for an election in a given city). Besides, it could be tested in a multi-lingual setting where it would generate suitable and multi-lingual news headlines for news articles covering the same topic. Another application could be altering written texts to include some metaphoricity. By defining a topic of an arbitrary text, the method could be adjusted to replace verbs and adjectives in the text with placeholders and, then, follow the method in filling them with content words from the interesting space while maximising the metaphoricity dimension.

11 Acknowledgements

This work has been supported by the Academy of Finland under grant 276897 (CLiC).

Bibliography

[1] Alnajjar, K., and Hämäläinen, M. 2018. A master-apprentice approach to automatic creation of culturally satirical movie titles. In *Proceedings of the 11th International Conference on Natural Language Generation*, 274–283. Tilburg, The Netherlands: Association for Computational Linguistics.

[2] Alnajjar, K.; Hämäläinen, M.; Chen, H.; and Toivonen, H. 2017. Expanding and

weighting stereotypical properties of human characters for linguistic creativity. In *Proceedings of the 8th International Conference on Computational Creativity (ICCC 2017)*, 25–32. Atlanta, United States: Georgia Institute of Technology.

[3] Alnajjar, K.; Hadaytullah, H.; and Toivonen, H. 2018. "Talent, Skill and Support." A method for automatic creation of slogans. In *Proceedings of the 9th International Conference on Computational Creativity (ICCC 2018)*, 88–95. Salamanca, Spain: Association for Computational Creativity.

[4] Alnajjar, K. 2018. The 12 Million Most Frequent English Grammatical Relations and their Frequencies. https://doi.org/10.5281/zenodo.1255800.

[5] Amabile, T. 2012. *Componential Theory of Creativity.* Number 12-096 in Working paper. Harvard Business School.

[6] Boush, D. M. 1993. How advertising slogans can prime evaluations of brand extensions. *Psychology and Marketing* 10(1):67–78.

[7] Burgers, C.; Konijn, E. A.; Steen, G. J.; and Iepsma, M. A. R. 2015. Making ads less complex, yet more creative and persuasive: The effects of conventional metaphors and irony in print advertising. *International Journal of Advertising* 34(3):515–532.

[8] Colton, S., and Wiggins, G. A. 2012. Computational creativity: The final frontier? In *Proceedings of the 20th European Conference on Artificial Intelligence*, ECAI'12, 21–26. Amsterdam, The Netherlands: IOS Press.

[9] Corbett, E. P. J. 1990. *Classical Rhetoric for the Modern Student.* Oxford University Press, 3rd ed. edition.

[10] Dahl, G. 2011. *Advertising for dummies.* John Wiley & Sons.

[11] Dahlén, M., and Rosengren, S. 2005. Brands affect slogans affect brands? competitive interference, brand equity and the brand-slogan link. *Journal of Brand Management* 12(3):151–164.

[12] De Smedt, T., and Daelemans, W. 2012. Pattern for Python. *Journal of Machine Learning Research* 13:2063–2067.

[13] Deb, K.; Pratap, A.; Agarwal, S.; and Meyarivan, T. 2002. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation* 6(2):182–197.

[14] Dowling, G. R., and Kabanoff, B. 1996. Computer-aided content analysis: What do 240 advertising slogans have in common? *Marketing Letters* 7(1):63–75.

[15] Dubovičienė, T., and Skorupa, P. 2014. The analysis of some stylistic features of English advertising slogans. *Žmogus Ir Žodis* 16(3):61–75.

[16] Evert, S. 2008. Corpora and collocations. In Lüdeling, A., and Kytö, M., eds., *Corpus Linguistics. An International Handbook*, volume 2. Berlin, Germany: Mouton de Gruyter. 1212–1248.

[17] Fortin, F.-A.; De Rainville, F.-M.; Gardner, M.-A.; Parizeau, M.; and Gagné, C. 2012. DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research* 13:2171–2175.

[18] Fuertes-Olivera, P. A.; Velasco-Sacristán, M.; Arribas-Baño, A.; and Samaniego-Fernández, E. 2001. Persuasion and advertising English: Metadiscourse in slogans and headlines. *Journal of Pragmatics* 33(8):1291–1307.

[19] Galvan, P.; Francisco, V.; Hervás, R.; Mandez, G.; and Gervás, P. 2016. Exploring the role of word associations in the construction of rhetorical figures. In *Proceedings of the Seventh International Conference on Computational Creativity (ICCC 2016)*. Paris, France: Sony CSL.

[20] Gatti, L.; Özbal, G.; Guerini, M.; Stock, O.; and Strapparava, C. 2015. Slogans are not forever: Adapting linguistic expressions to the news. In *Proceedings of the 24th International Conference on Artificial Intelligence*, 2452–2458. AAAI Press.

[21] Giora, R. 2003. *On our mind: Salience, context, and figurative language*. Oxford University Press.

[22] Glucksberg, S. 2001. *Understanding Figurative Language: From Metaphor to Idioms*. Oxford University Press.

[23] Hämäläinen, M. 2017. Semfi – Finnish semantic data for poem generation. https://doi.org/10.5281/zenodo.1137734.

[24] Hämäläinen, M. 2018. Harnessing NLG to create Finnish poetry automatically. In *Proceedings of the 9th International Conference on Computational Creativity (ICCC 2018)*, 9–15. Salamanca, Spain: Association for Computational Creativity.

[25] Harmon, S. 2015. FIGURE8: A novel system for generating and evaluating figurative language. In *Proceedings of the 6th International Conference on Computational Creativity (ICCC 2015)*, 71–77. Park City, Utah, United States: Brigham Young University.

[26] Honnibal, M., and Montani, I. 2017. spaCy 2: Natural Language Understanding with Bloom Embeddings, Convolutional Neural Networks and Incremental Parsing. *To appear.*

[27] Jordanous, A. 2012. A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation* 4(3):246–279.

[28] Klenosky, D. B., and Gitelson, R. E. 1997. Characteristics of effective tourism promotion slogans. *Annals of Tourism Research* 24(1):235–238.

[29] Kohli, C.; Suri, R.; and Thakor, M. 2002. Creating effective logos: Insights from theory and practice. *Business Horizons* 45(3):58–64.

[30] Labrador, B.; Ramón, N.; Alaiz-Moretón, H.; and Sanjurjo-González, H. 2014. Rhetorical structure and persuasive language in the subgenre of online advertisements. *English for Specific Purposes* 34:38–47.

[31] Lenzo, K. 1998. The CMU pronouncing dictionary. http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

[32] Loper, E., and Bird, S. 2002. NLTK: The natural language toolkit. In *Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*, ETMTNLP '02, 63–70. Stroudsburg, PA, United States: Association for Computational Linguistics.

[33] Manning, C. D.; Surdeanu, M.; Bauer, J.; Finkel, J.; Bethard, S. J.; and McClosky, D. 2014. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, 55–60.

[34] Mathur, L. K., and Mathur, I. 1995. The effect of advertising slogan changes on the market values of firms. *Journal of Advertising Research* 35(1):59–65.

[35] McGregor, S.; Agres, K.; Purver, M.; and Wiggins, G. 2015. From distributional semantics to conceptual spaces: A novel computational method for concept creation. *Journal of Artificial General Intelligence* 6(1):55–86.

[36] Mcquarrie, E. F., and Mick, D. G. 1996. Figures of rhetoric in advertising language. *Journal of Consumer Research, Inc. Source: Journal of Consumer Research* 22(4):424–438.

[37] Miller, D. W., and Toman, M. 2016. An analysis of rhetorical figures and other linguistic devices in corporation brand slogans. *Journal of Marketing Communications* 22(5):474–493.

[38] Nivre, J.; Hall, J.; Nilsson, J.; Chanev, A.; Eryigit, G.; Kübler, S.; Marinov, S.; and Marsi, E. 2007. Maltparser: A language-independent system for data-driven dependency parsing. *Natural Language Engineering* 13(2):95–135.

[39] Özbal, G.; Pighin, D.; and Strapparava, C. 2013. Brainsup: Brainstorming support for creative sentence generation. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 1446–1455. Sofia, Bulgaria: Association for Computational Linguistics.

[40] Reece, B. B.; Van den Bergh, B. G.; and Li, H. 1994. What makes a slogan memorable and who remembers it. *Journal of Current Issues & Research in Advertising* 16(2):41–57.

[41] Richards, J. I., and Curran, C. M. 2002. Oracles on "advertising": Searching for a definition. *Journal of Advertising* 31(2):63–77.

[42] Richards, I. A. 1936. *The Philosophy of Rhetoric.* London, United Kingdom: Oxford University Press.

[43] Searle, J. R. 1976. A classification of illocutionary acts. *Language in Society* 5(1):1–23.

[44] Strapparava, C.; Valitutti, A.; and Stock, O. 2007. Automatizing two creative functions for advertising. In Cardoso, A., and Wiggins, G., eds., *Proceedings of the 4th International Joint Workshop on Computational Creativity*, 99–108. London, United Kingdom: Goldsmiths, University of London.

[45] Tom, G., and Eves, A. 1999. The use of rhetorical devices in advertising. *Journal of Advertising Research* 39(4):39–43.

[46] Tomašic, P.; Papa, G.; and Znidaršic, M. 2014. Automated slogan production using a genetic algorithm. In *Student Workshop on Bioinspired Optimization Methods and their Applications - BIOMA 2014*, 55–66.

[47] Tomašic, P.; Papa, G.; and Żnidaršič, M. 2015. Using a genetic algorithm to produce slogans. *Informatica* 39(2):125.

[48] Tomašič, P.; Żnidaršič, M.; and Papa, G. 2014. Implementation of a slogan generator. In *Proceedings of the 5th International Conference on Computational Creativity (ICCC 2014)*, 340–343. Ljubljana, Slovenia: Josef Stefan Institute.

[49] Turney, P. D.; Neuman, Y.; Assaf, D.; and Cohen, Y. 2011. Literal and metaphorical sense identification through concrete and abstract context. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11, 680–690. Edinburgh, United Kingdom: Association for Computational Linguistics.

[50] Veale, T., and Li, G. 2012. Specifying viewpoint and information need with affective metaphors: A system demonstration of the metaphor magnet web app/service. In *Proceedings of the ACL 2012 System Demonstrations*, ACL '12, 7–12. Jeju Island, Korea: Association for Computational Linguistics.

[51] Veale, T., and Li, G. 2013. Creating similarity: Lateral thinking for vertical similarity judgments. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 660–670. Sofia, Bulgaria: Association for Computational Linguistics.

[52] Xiao, P., and Blat, J. 2013. Generating apt metaphor ideas for pictorial advertisements. In *Proceedings of the 4th International Conference on Computational Creativity (ICCC 2013)*, 8–15. Sydney, Australia: The University of Sydney.

[53] Xiao, P.; Alnajjar, K.; Granroth-Wilding, M.; Agres, K.; and Toivonen, H. 2016. Meta4meaning: Automatic metaphor interpretation using corpus-derived word associations. In *Proceedings of the 7th International Conference on Computational Creativity (ICCC 2016)*. Paris, France: Sony CSL.

[54] Żnidaršič, M.; Tomašic, P.; and Papa, G. 2015. Case-based slogan production. In Kendall-Morwick, J., ed., *In Proceedings of the ICCBR 2015 Workshops*, 123–130. Frankfurt, Germany: CEUR.

[55] Şimon, S., and Dejica-Carţiş, D. 2015. Speech acts in written advertisements: Identification, classification and analysis. *Procedia - Social and Behavioral Sciences* 192:234–239.

Appendix 1. Descriptions of Rhetorical Devices

Rhetorical Operation	Rhetorical Device	Brief description
Repetition (Sounds)	Rhyme	Repetition of syllables at the end of words
	Chime	Keywords in a phrase begin with identical consonants
	Assonance & alliteration	Three or more repetitions of a vowel or consonant
Repetition (Words)	Anaphora	Repetition of words at the beginning of phrases
	Epistrophe	Repetition of words at the end of phrases
	Epanalepsis	Repetition of a word toward the beginning and end of a phrase
	Anadiplosis	Repetition of a word toward the end of one phrase and the beginning of the next
Repetition (Phrase structure)	Parison	Marked parallelism between successive phrases; often involves the use of one or more embedded repeated words
Reversal (Syntax)	Antimetabole	Repetition of a pair of words in a phrase in reverse order
Reversal (Semantic)	Antithesis	Incorporation of binary opposites in a phrase
Substitution (Claim extremity)	Hyperbole	Exaggerated or extreme claim
Substitution (Assertive force)	Rhetorical question	Asking a question so as to make an assertion
	Epanorthosis	Making an assertion so as to call it into question
Substitution (Presence or Absence)	Ellipsis	A gap or omission that has to be completed
Substitution (Center or peripher)	Metonym	Use of a portion, or any associated element, to represent the whole
Destabilization (Similarity)	Metaphor	Substitution based on underlying resemblance
	Pun (general)	Substitution based on accidental similarity
Destabilization (Opposition)	Paradox	A self-contradictory, false, or impossible statement
	Irony	A statement that means the opposite of what is said

Table 15: Descriptions of rhetorical devices as given by Mcquarrie and Mick [36].

Appendix 2. Notations

Symbol	Description
T	The input target/tenor concept.
P	The input adjectival property to express.
V	List of metaphorical vehicles.
v	A metaphorical vehicle.
$\mathcal{E}$	A slogan.
t	An individual token in a given slogan.
c	Content words in a slogans.
f_{name}	A function for measuring a given linguistic characteristic denoted as $name$.
ζ	The text corpus.
ψ	A slogan corpus.
ξ	The language model.
ω	The semantic model.
δ	Database of expression skeletons.
γ	The repository of grammatical relations.
κ	Nouns and their adjectival properties.
ρ	Grammatical relations of a slogan.
φ	Phoneme/s obtained using CMU [31].
Υ	Universal semantic space.
I	The interesting semantic space containing valuable words.
s	A syntactical skeleton
G	The number of generations/iterations in the genetic algorithm.
μ	The population size in the genetic algorithm.
λ	The number of produced offspring in the genetic algorithm.
$Prob_c$	The probability of crossover in the genetic algorithm.
$Prob_m$	The probability of mutation in the genetic algorithm.
pos	A function returning individuals having all dimensions as positive, i.e. > 0.
max	A function returning individuals with the maximum value.
min	A function returning individuals with the minimum value.
μ_x	The mean.
SD	The standard deviation.
$sentiment$	A sentiment classifier provided by *Pattern* [12].

Table 16: A list of symbols defined in the thesis.

Appendix 3. Examples of Generated Metaphors

Tenor	Property	Apt	Strongly Related	Related	Random
book	valuable	purse	image	ginger	metal
painting	elegant	velvet	tuberose	aluminum	gps
car	elegant	scarf	tuberose	mahogany	mold
professor	smart	refrigerator	dolphin	weapon	pomfret
computer	creative	poet	performance	speech	bittersweet
professor	old	tractor	printer	beads	timber
politician	aggressive	bullying	wrestling	skateboarding	ambulance
chocolate	healthy	colon	herb	aorta	tantrism
museum	ancient	latin	brachiopod	universe	crocodile
love	beautiful	art	line	moonstone	deerskin

Table 17: 10 randomly selected examples of apt vehicles generated by the method using $V_{General}$, for the input tenor and property. The last three columns represent the other three vehicles selected during the evaluation.

Tenor	Property	Apt	Strongly Related	Related	Random
book	wise	father	judge	brother	marker
museum	scientific	scientist	computer	technologist	apartment
computer	powerful	king	tyrant	mogul	grief
politician	powerful	monster	emperor	thug	temple
professor	wise	king	father	politician	executive
coke	sweet	mother	friend	mistress	cinema
coke	dark	demon	terrorist	spy	travel
paper	scientific	scientist	computer	philosopher	hexachlorophene
professor	old	child	king	invalid	tendon
love	wild	cat	warrior	pirate	orator

Table 18: 10 randomly selected examples of apt vehicles generated by the method using V_{Human}, for the input tenor and property. The last three columns represent the other three vehicles selected during the evaluation.

Appendix 4. Examples of Generated Slogans

Input		**Output**	
Concept	**Property**	**Vehicle**	**Slogan**
coke	dark	demon	Resurrect Darkness. Fight More.
university	diverse	artist	Support Student. Study More.
love	beautiful	queen	Faith, Love And Faithfulness.
politician	powerful	monster	The Best Love Is Kind.
car	elegant	dancer	The Cars Of Stage.
computer	creative	poet	Win Big.
book	valuable	purse	The Best Item On Digitisation.
computer	powerful	supercomputer	Capability, Security And Reliability.
love	beautiful	queen	Happiness, Peace And Patience.
chocolate	sweet	candy	Fun, Food And Fruit.
painting	creative	literature	You Ca N't Sell The Fine Furniture.
politician	powerful	corporation	Filing The Public Defence.
painting	elegant	lady	Oils For A Traditional Painter.
coke	sweet	molasses	Ml For A Hot Heat.
politician	persuasive	orator	Excellent By Party. Speech By Talent.
coke	dark	basement	Fill Bottle. Live More.
politician	powerful	corporation	Filing The Governmental Request.
car	luxurious	marble	The Cars Of Portico.
politician	powerful	monster	Government Never Works Out Of Trust.
museum	scientific	scientist	Subjects For A Postdoctoral Fellowship.

Table 19: The top 20 slogans generated by the proposed method based on the judgements received on their overall suitability.

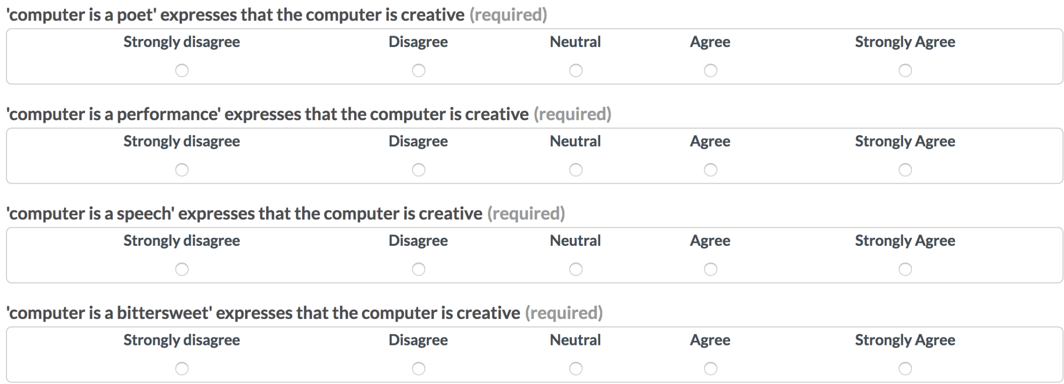
Input		**Output**	
Concept	**Property**	**Vehicle**	**Slogan**
coke	sweet	mother	Hear Your Candidates Do The Production.
painting	elegant	lady	Media For A Good Luck.
university	diverse	rhythm	Slower Than Offer.
chocolate	sweet	candy	Bill Going Soft.
politician	powerful	corporation	Filing The Exempt Certificate.
love	beautiful	art	ProductName. It Looks A Form Part.
paper	empty	jar	The Requirements Of Straight.
coke	sweet	molasses	Ml For A Woollen Manufacture.
love	wild	cat	Maximum People. Read Story.
professor	old	tractor	Department Formerly.
museum	ancient	queen	It Looks Mess.
university	diverse	rhythm	Slower Than Structure.
paper	white	shirt	Design Never Looks Out Of Sell.
book	wise	father	Find Your Children Do The Order.
paper	white	shirt	It Is Red.
paper	empty	jar	Piece Strip Story.
coke	dark	demon	Drink Bottle. Fall Faster.
computer	powerful	king	I're Operating It.
chocolate	healthy	colon	Between Terminal And Terminal Has Basket.
politician	powerful	monster	The Latest Creation Is Regulator.

Table 20: The least 20 slogans generated by the proposed method based on the judgements received on their overall suitability.

Input		**Output**	
Concept	**Property**	**Vehicle**	**Slogan**
museum	ancient	queen	It Is History.
paper	white	shirt	Number Never Gets Out Of Printer.
newspaper	international	dollar	Advert Never Appears Out Of Quarter.
car	exotic	olive	Supporting Students.
chocolate	healthy	child	The Fullest Support On Gift.
coke	dark	demon	Possess Power. Taste Better.
love	hungry	hyena	Live By Neighbour. Child By Kindness.
painting	majestic	eagle	Breathtaking The Rare Scenery.
museum	scientific	scientist	Find The Work To Date.
professor	smart	refrigerator	Philosophy, Physiology And Pharmacology.
politician	aggressive	predator	Politics For A Important Influence.
newspaper	international	dollar	Trying Times The Pm Over.
coke	dark	basement	Heave Sigh. Sit More.
car	elegant	scarf	Travel Available.
politician	persuasive	orator	Maker, Writer And Thinker.
professor	old	child	The Youngest Age Is Literature.
paper	scientific	scientist	Consultation, Collaboration And Publication.
book	wise	father	The Latest Puzzle On Faith.
love	hungry	hyena	Delighted By Affection. Peace By Dad.
university	diverse	artist	Between Title And Publisher Refers Reader.

Table 21: 20 randomly selected examples slogans generated by the proposed method.

Appendix 5. Examples of Crowdsourced Questionnaires



'computer is a poet' expresses that the computer is creative (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

'computer is a performance' expresses that the computer is creative (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

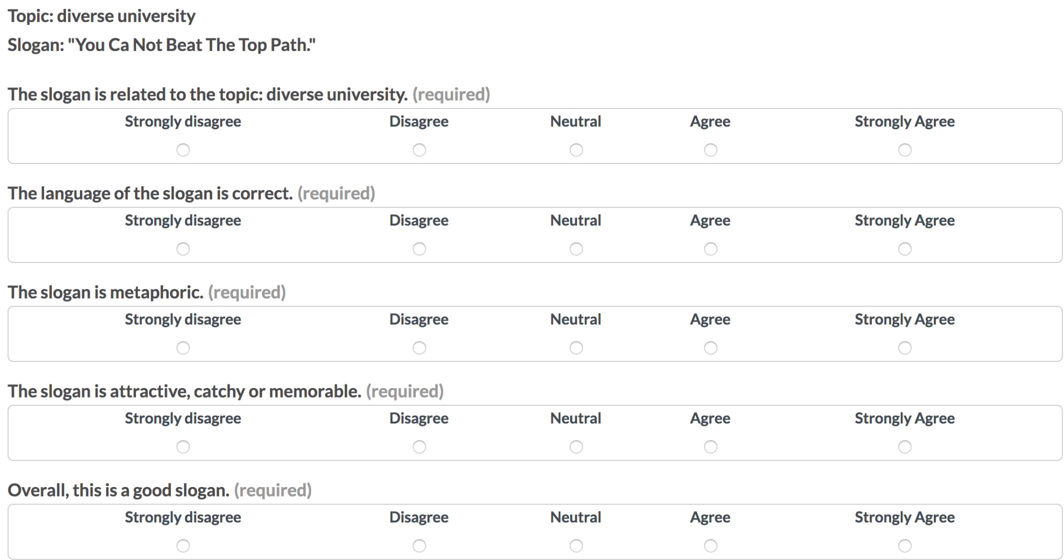
'computer is a speech' expresses that the computer is creative (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

'computer is a bittersweet' expresses that the computer is creative (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

Figure 8: An example of a crowdsourced questionnaire for evaluating vehicles for the tenor *computer* and adjectival property *creative*.



Topic: diverse university

Slogan: "You Ca Not Beat The Top Path."

The slogan is related to the topic: diverse university. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

The language of the slogan is correct. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

The slogan is metaphoric. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

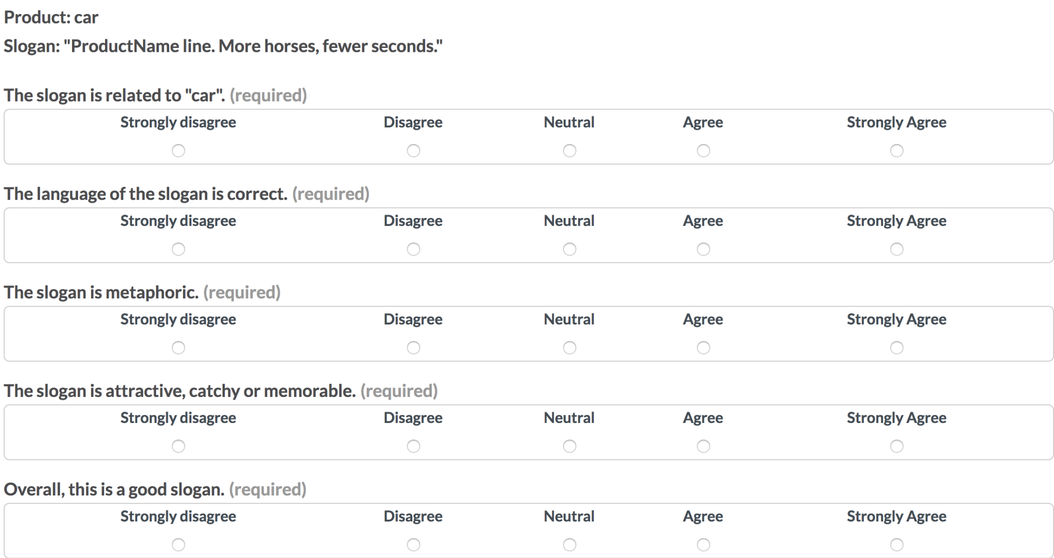
The slogan is attractive, catchy or memorable. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

Overall, this is a good slogan. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

Figure 9: An example of a crowdsourced questionnaire for evaluating a computer-made slogan



Product: car

Slogan: "ProductName line. More horses, fewer seconds."

The slogan is related to "car". (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
○	○	○	○	○

The language of the slogan is correct. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
○	○	○	○	○

The slogan is metaphoric. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
○	○	○	○	○

The slogan is attractive, catchy or memorable. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
○	○	○	○	○

Overall, this is a good slogan. (required)

Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
○	○	○	○	○

Figure 10: An example of a crowdsourced questionnaire for evaluating a human-made slogan