Humorous Similes

Abstract

Humorous descriptions are often couched in the form of a simile, whose flexible frame allows an author to yoke a topic to a perspective that is at once both incongruously different yet appropriately similar. Humorous similes exhibit all the commonly accepted hallmarks of verbal humour, from linguistic ambiguity to expectation violation and appropriate incongruity. But so too do non-humorous poetic similes, which exhibit an equal tendency for the ingenious and the incongruous. What then separates humorous similes from the broader class of creative similes, and can their signature characteristics, if any, be expressed via the presence or absence of specific formal, structural or semantic features? To address these questions, we describe the construction of a very large database of creative similes, and present the results of an initial empirical analysis upon this data-set. Our results are two-fold: humorous similes exhibit many of the same structural and semantic features that are considered characteristic of poetic similes, though none appears either necessary or sufficient to make a simile not just creative, but humorously creative; nonetheless, similes that employ either irony or ridicule (or both) are often explicitly marked with a marker of semantic imprecision such as "about". We go on to show that "about"-marked similes typically exhibit an identifiable affective signature that further telegraphs an author's humorous intent to the intended audience of the simile.

Keywords: Simile, Irony, Affect, Poetry, Computational Analysis

1. Introduction

The simile offers a highly productive form for the realization of memorable descriptions, so it is not surprising that similes are often used for humorous purposes, both in their stock, formulaic guise (see Taylor 1954; Norrick 1986) and in their more creative, exaggerated and ironic forms (e.g., see Fishlov 1992; Moon 2008). As the name suggests, a "simile" is a linguistic device for accentuating the similarities that exist between two otherwise different and dissimilar ideas, but when this balance of similarity and dissimilarity rises to the level of what Oring (2003) calls "appropriate incongruity" the result is often witty, insightful and just a little bit ridiculous. In this paper we argue that the simile form continues to be a frequent, varied and ingeniously crafted vehicle for expressing humorous intent, and we demonstrate its widespread use for humorous ends in everyday texts by compiling and analyzing a large corpus of creative comparisons from the world-wide-web. Our goal in this analysis is to attempt an empirical characterization of the structural and semantic properties of humorous similes in terms of the normative and non-normative properties that Fishlov (1992) has previously used to characterize another, related kind of creative similes – *poetic* similes. We shall describe an extensive semi-automatic process of corpus collection and analysis, using both manual and computational techniques, to identify those qualities that most consistently mark out a simile as potentially humorous.

A well-crafted humorous simile can deliver much the same semantic and pragmatic punch as a narrative joke. Consider example (1) from Raymond Chandler's *Farewell My Lovely* (1940):

(1) Even on Central Avenue, not the quietest dressed street in the world, he looked about as inconspicuous as a tarantula on a slice of angel food. (*Farewell, My Lovely*, 1940)

This comparison is clearly ironic, describing as it does a situation that is not just conspicuous, but startlingly eye-catching. This playful incongruity is further enhanced by Chandler's use of a wholly invented vehicle that exhibits its own internal surrealism, since one does not expect to find a deadly exotic spider on an innocent piece of cake. The

image is stark, immediate and easily grasped, and appears to obey Chandler's dictum (quoted in Speir 1981) that the primary purpose of simile is to "convey at once a simple visual image". Nonetheless, the visual incongruity of this simile, in which a large white person ("he" is "Moose Molloy") in a Black neighborhood is compared to a large black spider on a white cake, shows that Chandler is not so much striving for strict perceptual accuracy as he is for dramatic shock: he wants to make our skin crawl.

Chandler again taps into our sense of disgust in (2), emphasising social division by asking us to imagine an incongruous and unsavory food combination:

(2) I belonged in Idle Valley like a pearl onion on a banana split. (*The Long Goodbye*, 1954)

In (3), Chandler conjures a brutal image that seems at once both familiar (from iconic Hollywood movies) and shocking:

(3) The people who run that place are about as sympathetic as Georgia chain-gang guards. (*The Long Goodbye*, 1954)

This use of similes to convey a potent combination of perception and affective attitude has become a signature feature of the hard-boiled genre, as shown in the following example from Michael Chabon's Chandleresque *The Yiddish Policeman's Union* (Chabon 2007):

(4) He tugged in fits at the patchy remnant of his brown hair, or chased it with fingers back and forth across his pate like a pastry chef scattering flour on a marble slab.

The comparison of a vain balding man to an culinary expert (a pastry chef) does little to hide the cynicism in this elaborate description of a comb-over. We imagine instead the marble-smooth scalp, and the pathetic futility of its impermanent dusting of hair, Now consider this more typically poetic simile by Chabon, from his novel *Wonder Boys*:

(5) The whiskey tasted like bear steaks and river mud and the flesh of an oak-tree.

Though couched in the form of a simile, Chabon strives for a deeper metaphoricity here. The words "tasted like" are used to mean "evoked, when tasted" and we are not asked to contemplate the ridiculous idea of a whiskey that literally tastes like mud, wood and bear-meat. Extrapolating from Pollio (1997), it seems that where humorous similes engage

directly with the ridiculous, poetic similes stop short of this extreme and use metaphor to instead insulate us from the ridiculous consequences of taking a comparison at face value. But there is no structural or semantic cue here that asks us to see the metaphorical rather than the ridiculous; as an audience, guided by the provided co-text, we often collude with the author to choose the ridiculous interpretation that yields a humorous effect (see Veale 2004).

In *The Inimitable Jeeves*, Wodehouse (1975) offers the provocative imagery of (6):

(6) Aunt is calling to Aunt like mastodons bellowing across primeval swamps.

Once again the simile provides us with a striking mental visualization, one that allows us to fuse the imagery of aunt and mastodon to achieve extreme, and comic, effects. We might imagine the mastodons gossiping over cups of tea, or wearing fearsome horned-rimmed glasses.

Despite this pragmatic emphasis of humor on the ridiculous, humorous and poetic similes share many of the same structural and semantic properties. Each is a creative form that we recognize intuitively but which is hard (if not impossible) to formally define. Given this difficulty, Fishlov (1992) stops short of defining the essence of a poetic simile in outright terms, and prefers instead to define it as deliberate deviation from a set of non-poetic (and uncreative) norms. We intend to investigate here whether these norms provide an equally good basis for characterizing the ways that a humorous simile can harness the ridiculous to defy our expectations. In section 2 then, we consider the formal and semantic qualities that Fishlov argues are signal characteristics of creative similes, and in section 3 ask whether Fishlov's framework for the analysis of poetic similes is equally applicable to the analysis of humorous similes. A fuller exploration of creative similes requires a large corpus of many different examples from many different sources, rather than a cherry-picked collection of literary gems. To avoid confirmation bias, the corpus should be automatically compiled, from as large an author-base as possible. We therefore describe the acquisition and annotation of such a corpus from the world-wide-web in section 4, and provide an affect-oriented analysis of this corpus in section 5. We summarize our findings in section 6, where we consider the degree of empirical support for our conclusions

2. Humorous Similes as Poetic Comparisons

Fishlov (1992) argues that a simile is poetic to the extent that it diverges from the norms of the non-poetic simile form. In Fishlov's view, poetic similes (which he designates PS) can depart from non-poetic (or NPS) norms on one or more of eight different structural and semantic dimensions. Fishlov identifies these dimensions as follows: (i) **order** – conventional similes observe the sequence *Topic [comparator] Vehicle*, while poetic similes often play with this order by e.g., introducing the vehicle first; (ii) **length** – NPS similes are conventionally short and punchy (e.g., see Taylor 1954), while PS similes often contain a surfeit of picture-building detail; (iii) **explicitness** – highly conventionalized similes tend to be unambiguous about the features that the vehicle lends to the topic, while more creative PS similes encourage ambiguity and allow for a range of possible interpretations; (iv) **literality** – the features highlighted by similes conventionally apply directly to the topic, without the need for figurative analysis of their potential meaning, while poetic similes often demand precisely this kind of additional metaphoric accommodation; (v) **salience** – conventional similes employ vehicles whose transferred features are highly salient or even stereotypical (Ortony 1979), whereas creative similes often cast vehicles against type; (vi) **familiarity** – conventional similes illuminate a topic that may be unfamiliar to an audience by using a vehicle that is both familiar and well-understood, while creative similes often use unrealistic or bizarre scenarios to view everyday life from a fresh perspective; (vii) **connotation** – conventional similes use words whose connotative tone is suited to the topic, and unlike humorous similes, do not use words that evoke layers of gratuitous meaning and innuendo; and (viii) **distinctiveness** – we expect the topic and vehicle of a simile to be quite different ideas, yet creative similes may, in some rare cases, employ a vehicle category to which the topic obviously belongs.

Fishlov effectively advocates a radial category structure for similes, in the vein of Lakoff (1987), in which NPS norms occupy a central position and more poetic cases reside further out, at a radial distance defined by their differences from the prototypical norm. Examples (7) – (14) each exemplify a PS departure from a different NPS norm.

(7) **Order**: As when a torrent, swell'd with wintry rains, Pours from the mountains o'er the delug'd plains, … Fierce Ajax thus overwhelms the yielding throng [*Homer's Illiad, Book XI.* English Translation by (Pope 1813:458)]

(8) **Length**: You mean as cunning as a plan devised by a fox who is so cunning that he has been elected Professor of Cunning at Oxford University? [from the TV show *Blackadder Goes Forth* (Lloyd et al. 1999)]

(9) **Explicitness**: It tasted like a high fog strained through cotton wool. [*The Long Goodbye* (Chandler, 1954)]

(10) **Literality**: His research is about as ground-breaking as a foam jackhammer.

(11) **Salience**: The voice got as cool as a cafeteria dinner. [*Farewell, My Lovely* (Chandler, 1940)]

(12) **Familiarity**: when Boomer danced he looked like a monkey on roller skates juggling razor blades in a hurricane. [*Skinny Legs And All*, Robbins (1990)]

(13) **Connotation**: Her face was shining like the seat of a bus-driver's trousers. [*Bertie Wooster Sees It Through*, Woodhouse, (1955)]

(14) **Distinctiveness**: The grave opens up before me like a big hole in the ground. [from the television show *Blackadder* (Lloyd et al. 1999)]

Fishlov observes that PS similes typically diverge from NPS norms on a number of dimensions simultaneously, combining excessive length or a twisted ordering with at least one semantic or pragmatic deviation from the norm. Fishlov's framework is thus very much in the mold of both Giora (2002) and Hanks (2004). Giora's notion of *optimal innovation* requires that, to be effective, the novelty of a creative formulation must be rooted in the familiar and the understandable, while Hanks argues that most creative expressions are subtle *exploitations* of more familiar forms. Irony is perhaps the most common exploitation of NPS norms for wringing humour from similes, as in (10). Such similes exhibit a degree of pretence or pragmatic insincerity on the part of the speaker (Clark and Gerrig 1984) as well as an allusion to a salient expectation that is violated by the utterance (Kumon-Nakamura, Glucksberg and Brown 1995). This ironic allusion often echoes an assumption that the speaker wishes to undermine (Sperber and Wilson 1992), especially if the simile forms part of an adversarial exchange (Veale et al. 2006).

The Homeric simile in (7) shows the poetic value of a *vehicle-before-topic* ordering, while (8) to (4) illustrate the humorous potential in violating other NPS norms. While these latter criteria are clearly applicable to humorous similes, we now consider what they overlook.

3. The Affective Dimension of Humorous Similes

Though incongruity is implicit in the violation of any norm, Fishlov's framework only hints at a consideration that is so often given centre-stage in humor research, logical incongruity (e.g., see Suls 1972; Raskin 1985; Attardo and Raskin 1991; Attardo, 1997; Ritchie 1999; Attardo et al. 2002; Oring 2003; Veale 2004). Incongruity, after all, is a deviation *par excellance* from the norm, both semantically and pragmatically. Humorous similes more often conjure images of the ridiculous than the sublime, and though Fishlov's connotative dimension broadly encompasses this kind of incongruity, affect is so important to the working of humorous similes that it deserves a distinct dimension of its own. To appreciate the affective difference between similes that are obviously humorous and obviously poetic, it is useful to consider a text in which both kinds sit side by side. Consider (15), an extract from the Leonard Cohen song *Bird on a wire*:

(15) Like a bird on the wire,
Like a drunk in a midnight choir,
I have tried, in my way, to be free.
Like a worm on a hook,
Like a knight from some old-fashioned book,
I have saved all my ribbons for thee.
(Cohen, 1994, page 144)

Some vehicles appear more familiar than others here, though each is given a poetic cast by the lyricism of Cohen's text. For instance, "a knight from some old-fashioned book" is at once very familiar and evocative, and poetic too by virtue of the linguistic trickery that allows Cohen to shift the modifier "old-fashioned" from "knight" to "book" so that the rhythm and rhyme of the lyric is maintained. In the course of this fragment, the singer is compared to two kinds of animal, a "bird" and a "worm", and two kinds of person, a "knight" and a "drunk", and though the comparison of the singer to a "worm" should appear the most incongruous, and the most disrespectful, it is the comparison to a "drunk" that generates the most humor. The reasons for this are two-fold. First, the image of a drunk in a choir is both novel and internally incongruous, since we expect choralists to be committed volunteers; the image of a worm on a hook is neither. Second, and perhaps most importantly, the image is not just striking but ridiculous: we imagine the drunk singing with gusto while earning the enmity of his frustrated colleagues,

and just as we would be entertained by the foolishness of the drunk in real-life, were we to witness such an event, we find the image constructed here foolish and entertaining. It is not just the incongruity, but the affective tone of the incongruity, that prompts us to laugh. In the following sections we shall empirically investigate whether humorous similes do indeed have a distinctive affective signature.

4. Compiling A Comprehensive Database of Similes

To understand the pivotal role of affect in humorous similes, we shall need to gather and annotate a very large corpus of similes as they are used in many different types and genres of text. The richest source of these similes is the collected texts of the world wide web. We look first to the most conventional, and most prevalent, uses of similes, before turning our gaze to more creative and less frequent uses.

4.1. Harvesting Simple Similes with One-Word Vehicles

To compile a collection of conventional similes, one can look to authoritative sources such as printed dictionaries, or exploit the syntactic frame of the *as*-simile to identify matching instances in large text corpora. Norrick (1986), for instance, uses the former approach, and bases his analysis on 366 similes listed in the 1970 edition of *The Oxford Dictionary of Proverbs*. Moon (2008) uses a hybrid approach, and compiles a collection of 377 similes from multiple sources, one of which is the *Bank of English* corpus. But the pervasiveness and ease of use of the simile form means that one is likely to find a greater diversity in the collected texts of the world-wide-web (Roncero et al. 2006).

The syntactic marking of similes means that most can be harvested automatically from the web, using a simple process of pattern-matching. Thus, when we pose the query “as * as *” to the Google search-engine (*www.google.com*), the wildcard elements are bound by Google to the corresponding elements of the comparison form in (16):

(16) “as * as *”

Presented with queries in the form of (16), the search-engine returns a large number of snippets from online documents that contain matching phrases, such as “as hot as an oven” or “as strong as an ox”. In these snippets, we are likely to see the same combination of ground and vehi-

cle recur in many different contexts. This combination of ground and vehicle is the semantic core of a simile, the part that transcends context to be reused in the description of many different topics. The relationship of this core combination to the topic, will in many cases, be entirely contingent and subjective; most similes are used, after all, to communicate information about a topic that is not fully understood or appreciated by an audience, and so for purposes of corpus construction, the topic has very little bearing on the semantics of the simile. For instance, the simile “my boss is as cunning as a fox” tells us nothing at all about bosses in general, but does tell us that foxes are either stereotypically cunning (if the simile is non-ironically *straight*) or stereotypically naïve (if the simile is ironic). We are primarily interested therefore in the collection of simile *types* – the context-transcending reusable combination of a specific ground with a specific vehicle – rather than of simile *instances* – the contextually-tied application of a ground and vehicle to a specific topic. When one considers that pronouns are often used as the topic of a simile (as in (1), where “he” refers to “Moose Molloy”), or that topics may not be fully realized in the text surrounding a simile (if the topic is already known to the audience), a large-scale semantic analysis can only meaningfully focus on the pairing of ground and vehicle in types rather than instances.

A web search-engine is designed to return a diversity of different documents rather than a diversity of specific text matches, and so a query like (16) may find many simile instances that correspond to a small set of the most common simile types. To ensure that we acquire the widest range of simile types with the widest range of adjectival grounds, we need to seed our queries with specific adjectives. For example, to ensure that we find similes for *strength*, we need to use the queries “as *strong* as *” and “as *weak* as *”. To automate the harvesting process, we use the lexical resource WordNet (see Fellbaum 1998) as an inventory of antonymous adjective pairs, such as “strong” and “weak”, since these often define the gradable properties for which similes are used to provide specific values. In all, we generate over 2000 queries with the form of (16), in which the ground position is successively bound to a different adjective from the WordNet inventory.

It is impractical to consider every single document indexed by Google as a match for these 2000 queries. Rather, we consider just the first 200 snippets returned for each query, allowing us to harvest a corpus of simile types by taking a wide-ranging series of different core-samples from

across the full breadth of the web. While the core-sample for each adjective is just 200 snippets deep, this is sufficient for a frequency analysis to reveal the most culturally entrenched English similes. For instance, in the query "as *strong* as *", the wildcard * matches "horse" 27 times, "bull" 19 times, "gorilla" 12 times, and "Viking" just once.

These simile types are restricted to grounds that appear on our initial 2000-adjective list, which is extensive but not exhaustive. To go beyond this list, we use the extracted vehicles from this first harvesting sweep (such as "horse", "gorilla" and "Viking" above) to seed a second set of queries in which the vehicle rather than the ground is specified, such as "as * as a *bull*" and "as * as a *gorilla*". For instance, the query "as * as a *Viking*" allows us to identify the ground "blonde" that is not on our original list of antonymous adjectives. This second phase is especially useful for finding similes with rare vehicles such as "Viking": if at least one simile type involving such a vehicle is found in the first phase, then additional types with this vehicle are targeted during the second phase.

When we consider only those simile instances with a single-word vehicle, as listed in a conventional lexical resource like WordNet, the above harvesting phases yield 74,704 instances of the "as * as *" pattern, 42,618 of which are unique. In all, these instances relate 3769 different adjectival grounds to 9286 different noun vehicles. However, while each of these instances is a legitimate instance of a comparison, not all qualify as similes. As defined by Ortony (1979), the difference between comparisons and similes is best characterized in terms of salience: a simile uses a vehicle for which a given ground property is especially salient to highlight this property in a topic (this is Fishlov's salience dimension). Simple comparisons, on the other hand, merely point out the commonalities between two things, regardless of whether those properties are highly salient in the vehicle. If a doctor states that a tumour is "as big as a tennis-ball", this may well be cause for alarm, but it is not a simile, since bigness is not a salient property of tennis-balls.

Since there is no automatic way of separating similes from simple comparisons, this separation must be performed by hand. Human judges are therefore used to annotate all those instances where the ground is obviously a salient property of the vehicle (the bona-fide or *straight* cases) or where a property that is diametrically opposed to the ground is salient of the vehicle (the ironic cases). The extensive grey area between these positions – where the ground is neither strongly associated with, nor strongly opposed to, the vehicle – is not always clear cut, and

instances like "as cuddly as a bear" might fall into either category in one context or another. The human judges are asked to perform a conservative separation, discarding those instances that might lean one way or another in different contexts. Those instances that are not discarded are annotated at this time as either straight or ironic. In all, 30,991 instances are identified as straight (non-ironic) similes; these instances provide 12,259 unique simile types, that is, unique pairings of a ground property to a vehicle. A smaller but significant body of 4685 instances are annotated as ironic, such as "as hairy as a bowling-ball", providing 2798 unique ironic types.

4.2. Harvesting More Complex Affect-Rich Similes

Similes are hedged assertions, in which a topic is merely stated to be approximately similar to, and not absolutely identical to, a given vehicle (recall that this is Fishlov's *distinctiveness* dimension). Indeed, some similes are doubly-hedged, as though to indicate that the case for similarity is even more approximate. We see double-hedging in (1) and (3), where Chandler uses "about" to mark the tongue-in-cheek nature of his comparisons. Interestingly, Moon (2008) states that the "about" marker always signals the use of irony in similes, but our analysis in section 5 does not bear out so strong a claim. Rather, we argue that "about" telegraphs an author's intention to use an inventive vehicle which merely exhibits an inexact ballpark similarity to the topic. It other words, "about" in a humorous simile means "not precisely", and often provides the implicit negation that Giora (1995) argues is intrinsic to irony, signaling the use of a vehicle that is, in the terminology of Oring (2003), *spurious*. Of course, ironic similes are highly spurious, and are thus likely to be prefixed by the "about" marker to signal this fact to an audience, but there are other strategies for achieving humor through similes – such as exaggeration, pointed insult, and absurdity – and these too may well be prefixed by "about". To settle the issue, we acquire a large corpus of "about" similes using a query tailored to the task:

(17) "about as * as *"

We therefore re-run the harvesting processes with the query form in (17), and extract all syntactically well-formed vehicles, whether they comprise one word or many, since excessive vehicle length is one of the structural properties of creative similes identified by Fishlov. The re-

trieved instances thus run the gamut from the short and punchy to the long and overwrought, as typified by (18) and (19) respectively.

(18) about as pervasive as air.

(19) about as difficult as finding work as a school teacher after a child-abuse conviction.

In all, this second sweep of the harvester yields 45,021 instances of the "about" pattern in (17). Most of these instances occur just once overall, and this second harvesting sweep yields almost as many unique types (38,294) as instances, suggesting that 85% of these instances are creative one-offs. When hand-annotated for the salience profile we expect from similes, we find that 20,299 of these types (53%) are more than mere comparisons, and use vehicles for which the ground is either very salient or ironically opposed.

Interestingly, just 14% of these 20,299 "about" simile types use a vehicle with a single content-word, and a mere 3% of these (i.e., 676 types) are also found in the original harvesting process of simple similes using query (16). In other words, the overlap in simile types found using both harvesting processes – simple similes using query (16) and "about" similes using query (17) – is negligible. Clearly, the addition of an "about" marker causes the second web sweep to harvest an almost completely different set of similes. From a Fishlovian perspective then, we see a clear quantitative and qualitative separation between similes that are marked with "about" from more conventional similes. The "about" similes are typically longer, with a mean size of three words per vehicle, excluding initial determiners. They are also more heavily inclined toward the ironic. Hand-annotating for straight or ironic descriptions, we find that only 4797 unique simile-types (or just 24%) employ a vehicle for which the ground is both salient and apt, while 15,502 simile-types (76%) are ironic, as in (20):

(20) about as modern as a top-hatted chimneysweep.

The "about" form thus seems to be syntactic framing that allows an author to telegraph an attempt to mint an unconventional and creative simile, of the kind that Fishlov designates *PS* for poetic. The strong propensity for irony with this form (76%) suggests that the double-hedging is intended to alert the audience to the possibility of irony, and to minimize the risk that the author's creative intent is misunderstood.

4.3. A Snapshot of Two Corpora

We have used the Google search engine to harvest two very different collections of similes from the web. The first corpus, using the query in (16), is biased toward the simple, one-word, conventionally non-poetic similes that Fishlov has designated *NPS*. The second corpus, gathered using the "about"-marked query in (17), is biased toward the longer and more original similes that Fishlov has designated *PS*. The overlap between both collections is minimal, suggesting that the "about" hedge is a good indicator of whether a simile's author strives for creative impact.

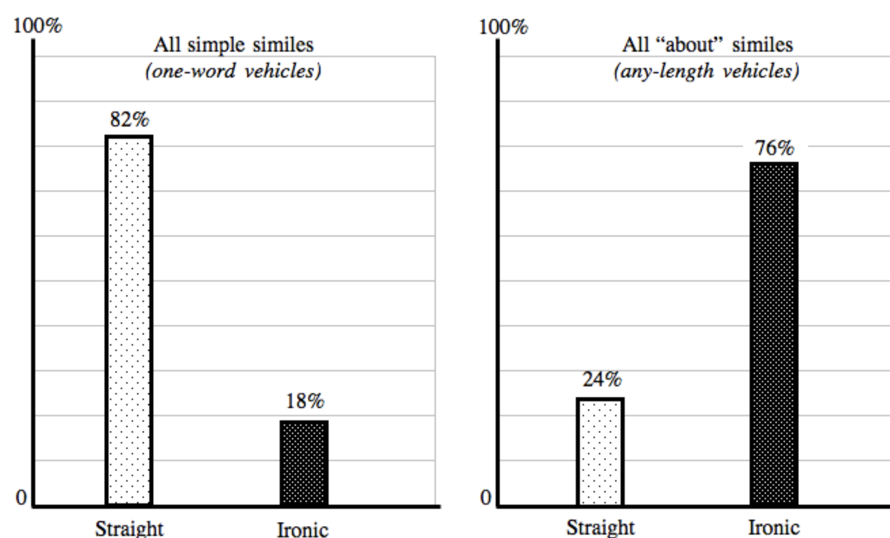

Figure 1. Relative frequency of straight vs. ironic similes in the two annotated web-corpora that are collected using queries (16) and (17) respectively.

Figure 1 shows that the presence of "about" signals a dramatic difference in the character of a simile. As revealed in the almost complete reversal of distributions between the graphs of Figure 1, "about" has the power to humorously up-end the semantic interpretation process. But as we shall now see, "about" alone does not make a simile ironic; it merely biases our comprehension processes toward a more humorous interpretation.

5. Affective Differences: Comparing Both Corpora

Though there is just a 3% overlap between the longer "about" similes and the shorter, more conventional figures of speech, we find that 62% of the "about" similes use at least one stock image drawn from the inventory of conventional vehicles. The longer similes do not use these stereotypes in isolation, or even to exemplify the same grounds, but combine them in novel ways to create strikingly new images (recall Tom Robbins' humorous combination of commonplace images in (9)) For instance, (21) and (22) are drawn from our first corpus of simple similes, while (23) is drawn from our second corpus of "about" similes.

(21) as quiet as a cat

(22) as noisy as a blender

(23) about as soothing as a cat in a blender

In a substantial number of the web-harvested "about" similes – 30% – the vehicle is a composite structure in which two or more concepts are linked via a preposition, as in (23). The combination of (23) employs two stock images with contrary properties – the stealthy cat and the loud blender – to evoke a visceral feeling of unease and disgust that stands in ironic opposition to the stereotype of calm relaxation that the simile initially promises. Notice how the simile cleverly plays each stock image against type: the cat, which might be considered soothing in normal circumstances, is placed in a cruel situation that prompts us to feel its suffering; and the blender, which is stereotypically loud and jarring, is ironically put forward as an exemplar of the very opposite. So while the longer "about" similes like (23) achieve more imaginative and creative effects than their conventionalized brethren, they are not completely distinct. They frequently draw upon the same conventional imagery, but in combinations that are designed to subvert stereotypical properties and create a heightened sense of perception and affect.

Complementing this stock imagery, we find that 12% of "about" similes make use of well known names from the current pop-cultural and political climate. The most common entities comprise a roll-call of topical hate-figures, from "Karl Rove" and "George Bush", to "Paris Hilton" and "Michael Moore". These entities are evocative enough to serve as complete vehicles in their own right, yet occur just as frequently in combination with other scene-setting concepts:

(24) about as lost as Paris Hilton in a library

(25) about as frustrated as Stevie Wonder in an Easter egg hunt

While it is hard to imagine (24) and (25) having any linguistic currency in a hundred years time, they do succeed in putting a topical face on some well-worn humor vehicles – "dumb" blondes and blind men – while exploiting the audience's feelings about these targets to sharpen their humorous edge.

5.1. Estimating The Positive / Negative Affect of Similes

A critical attitude is typical of irony, and creative "about" similes should be no different in this respect than simple similes with short, single-word vehicles. However, while some adjectives are uniformly critical in any context, such as "dull", "unattractive" and "stupid", most adjectives (such as "fragile", "tough" and "controversial") occupy a usage-sensitive middle ground between clearly-positive and clearly-negative. Lacking specific knowledge of a speaker's views on a topic, or indeed of the topic itself, the quantification of a simile's positive or negative affect is too subjective to be meaningfully performed by a small group of human annotators. To achieve as much consistency as possible in the rating of attitudes, we turn to Whissell's (1989) *dictionary of affect*, an inventory of over 8000 English words with pleasantness scores that are statistically derived from human ratings. These scores range from 1.0 (most unpleasant) to 3.0 (most pleasant), with a mean of 1.84 and a standard deviation of 0.44. For our current purposes, we assume the ground of a simile to be negative if its pleasantness is at least one standard deviation below the mean (≤ 1.36), and positive if its pleasantness is at least one standard deviation above the mean (≥ 2.28).

According to these criteria, a sub-set of 7256 "about" similes have clearly positive or negative grounds. This sub-set represents an almost perfect statistical sampling of the larger "about" corpus, with much the same distribution of straight versus ironic examples (21% / 79% versus 24% / 76%), so we can safely use this set to characterize the balance of affective attitudes in different kinds of simile. Figure 2 illustrates the breakdown of similes by straight/ironic, positive/negative and simple/"about" criteria. In the most conventional straight similes, we observe that a positive attitude is conveyed twice as often as a negative attitude. In contrast, simple ironic similes (with one word vehicles) con-

vey a negative attitude six times more often than a positive attitude. Turning to the more creative "about" similes (on the right), we see that straight "about" similes communicate a negative attitude a little more often than a positive attitude, but that ironic "about" similes carry a negative affect in almost 9 out of 10 cases. Simple similes are thus more likely to impart a positive view of a topic, while longer similes that carry the "about" marker are more likely overall (whether straight or ironic) to impart a negative view of a topic.

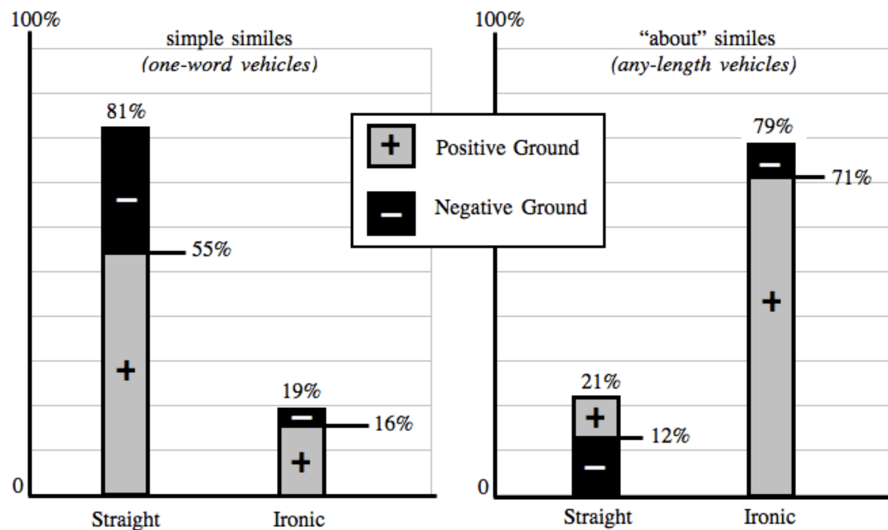

Figure 2. Distribution of positive and negative grounds across straight and ironic instances of simple (one-word vehicle) and "about" similes.

This difference is exacerbated by the strong preference for irony with the "about" form. Recall from the previous section that 76% of "about" similes (by type) are ironic, while just 18% of the shorter, more conventional similes are ironic. Overall then, 83% of "about" similes impart a negative view of a topic, since 12% of "about" similes are non-ironic with a negative ground, and 71% ironically use a positive ground to impart a negative property.

These numbers suggest not just that irony is widely used in simile, but they also begin to explain why it is used. Figure 2 (left) shows that nega-

tivity is under-represented in simple similes, and that straight conventional similes communicate a positive description more than twice as often as a negative description (55% versus 26%). Irony provides a necessary corrective to this imbalance, allowing negative descriptions to be crafted from positive grounds. Simple ironic similes thus have a predominantly negative meaning (16% versus 3%, Figure 2 left). Figure 2 (right) shows that "about" similes more than correct the remaining imbalance by choosing to employ their increased length and ingenuity in the service of negativity and ridicule.

6. Discussion and Conclusions

In *A Christmas Carol*, Charles Dickens suggests that conventional similes are a kind of linguistic *hand-me-down*, a rich source of cultural wisdom that is inherited by the speakers of a language:

> Old Marley was as dead as a door-nail. Mind! I don't mean to say that I know, of my own knowledge, what there is particularly dead about a door-nail. I might have been inclined, myself, to regard a coffin-nail as the deadest piece of ironmongery in the trade. But the wisdom of our ancestors is in the simile.
>
> (Charles Dickens, *A Christmas Carol*, 1984 [1843], page 1)

This cultural wisdom is entrenched in our language in highly conventionalized forms (Taylor 1954), and a simile marks itself out as creative to the extent that it deviates from these norms and conventions (Fishlov 1992). Humorous similes additionally mark themselves out by embracing not just the figurative but the ridiculous, but there is a very fine line indeed between these two notions. The appreciation of humorous intent in a simile needs not only a Fishlovian understanding of the simile's structural and semantic qualities, but a pragmatic understanding of which side of this line an author wishes to occupy. Consider the example in (26), harvested as part of our corpus of "about" similes from the web:

(26) He is about as tough as a marshmallow cardigan.

Though physically possible, the vehicle here is clearly fantastical, and crosses the line into humor by appealing to our sense of the ridiculous. The simile thus operates on three different levels: on the first level, the "marshmallow cardigan" evokes the properties of softness and weakness that the author wishes to ascribe to the topic; on the second level, the

ironic distance between these properties and the stated ground suggests that the ascription is contrary to expectation, and perhaps contradicts the belief system of the topic himself (who may simply “act tough”); and finally, the ridiculousness of the vehicle accentuates the irony by implicitly communicating the belief “It is not just wrong, but ridiculous, to believe such a person is tough”. Though the combination is clearly incongruous, defying our conventions of how clothing is made and food-stuffs are used, it is not inherently ridiculous, but is made to seem so by our appreciation of the author’s playful intentions. To see why, consider a comparable description in (27) from poet Vladimir Mayakovsky:

(27) If you wish,
I shall grow irreproachably tender,
not a man, but a cloud in trousers!
(*The Cloud in Trousers*, translated from Russian in Blake, 1975)

It is hard to find a purely semantic reason why the combination “marshmallow cardigan” is laughable and ridiculous while the combination “a cloud in trousers” is not. Both combine a common article of clothing with an object that stereotypically evokes notions of softness and lightness. Indeed, testifying to the empirical observation in the previous section that creative comparisons frequently employ stock imagery from more formulaic similes, albeit in novel combinations, our corpus of simple similes from section 4.1 contains four relevant comparisons: “as soft as a marshmallow”, “as soft as a cloud”, “as light as a marshmallow” and “as light as a cloud”. This corpus reveals further similarities between clouds and cardigans in the attested similes “as comfy as a cloud”, “as comfortable as a cloud”, “as comfy as a cardigan” and “as comfy as a cloud”, strongly suggesting that the key difference between “a marshmallow cardigan” and “a cloud in trousers” is not a semantic or conceptual one, but a pragmatic one.

There are suggestive differences in form between (26) and (27), not least the fact that (26) is a simile and (27) uses a metaphor. Readers approach Mayakovsky’s poem about an angry, spurned lover with a mind-set that befits a serious work about a deeply emotive subject, and are predisposed to see Mayakovsky’s combination as a poetic metaphor that acts as a place-holder for a complex emotional category (see Glucksberg, 2001), one that can give rise to an interesting and varied body of inferential mappings (see Lakoff and Johnson, 1980). As such, readers see past the physical composition and judge the metaphor more in terms of its high-minded target than in terms of its superficially ri-

diculous source. In contrast, readers recognize that "marshmallow cardigan" is not a metaphor for something else, but a physical combination that must be judged on its own terms. It thus evokes a disposable image with an immediate, scornful affect, but with no lasting metaphoric value. Though metaphors can also be humorous, the sensibility of a metaphoric vehicle is judged for its role in securing a cohesive integration of source imagery and target ideas (see Pollio 1997), while the meaning of a simile vehicle must, for the most part, stand alone. This relative lack of integration between a simile's vehicle and topic allows a greater disconnect of ideas and connotations, greater scope for the generation of ridiculous imagery, and greater scope for the generation of humor.

We have presented substantial empirical evidence that the "about" form acts as a scaffolding structure for humorous similes, priming an audience to view comparisons with positive grounds as ironically critical and comparisons with negative grounds as plainly critical. We employ the term *scaffolding* in the sense of Veale and Keane (1992), to mean a structure that allows immediate but superficial interpretation of a figurative utterance, and on which a deeper and more insightful interpretation can gradually be elaborated. In other words, the "about" form allows an audience to quickly construct a basic and mostly accurate interpretation of a speaker's intent without having to fully understand the meaning of the vehicle or having to resolve any incongruities the vehicle may contain. All that is required is that the audience be able to determine the intended affect – positive or negative – of the simile's ground: if correctly ascertained as positive, then the simile has close to a 90% chance of being ironic and critical; if ascertained as negative, the simile has just a 40% of being ironic and is 60% likely to mean what it overtly says.

Roncero et al. (2006) note that similes found on the internet are far more likely than the equivalent metaphors to be accompanied by an explicit explanation, suggesting that simile authors feel a need to cue readers as to the proper interpretation of their creative efforts. Explanations rob jokes of their potency, so we can expect humorous similes to eschew explicit explanations. In particular, ironic comparisons would be utterly undermined if accompanied by an explanation, since as Grice (1978:125) notes, "to announce [irony] as a pretence would be to spoil the effect". Nonetheless, irony always runs the risk of being misdiagnosed (Sperber and Wilson 1992), and so requires that care is exercised in its use. Grice (1978:125) further notes that when "speaking ironically … a tone suitable to such a feeling or attitude seems to be mandatory".

So when ironic comparisons are creatively minted on the fly, in conversationally-styled texts, it is intuitive to suppose that some form of lexicalized support structure will often be used in place of an ironical tone, to subtly direct the audience toward the desired meaning. The "about" marker is the textual equivalent of a raised eyebrow, a cue that signals a playfulness on the part of the author and one that licenses the audience to seek out a more humorous interpretation when one is available.

Yet the presence of "about" does not make a simile humorous, nor does its absence undo any potential a simile may have for humor. Though we can identify structural and semantic features of similes that contribute to their humorousness, we cannot identify structural or semantic features that are always sufficient to make a simile humorous, or for that matter, to make a simile poetic. Humor is not semantically or structurally determined, but arises from the pragmatic effects of an utterance's use in context (see Veale 2004). Nonetheless, structural properties – like the presence of "about" – can encourage an audience to collude with the author in constructing a humorous interpretation. The "about" form is unlikely to be the only construction that supports and primes a humorous interpretation in this way, though it does seem to be one of the simplest and most direct, at least for similes. Further analysis of our simile corpora is thus warranted, on the degree to which markers like "about" signal not just ironic incongruity between vehicle and ground, but humorous incongruity within the vehicle itself. This analysis will enable us to tease apart the key distinctions between the absurd (that which is logically impossible) and the ridiculous (that which is possible but inadvisable, wrong-headed or just stupid), and quantify the degree to which each is employed for humor-generation in similes.

Finally, we conclude by noting that our web-corpus of conventional similes is now available to browse on the web (at *http://afflatus.ucd.ie/* under *Sardonicus*), while a computational metaphor-interpretation system built using these similes as a knowledge-base, named *Aristotle*, is also publicly available to use online (at *http://afflatus.ucd.ie/aristotle/*).

References

Attardo, Salvatore & Victor Raskin
1991 Script theory revis(it)ed: joke similarity and joke representational model. *Humor: International Journal of Humor Research,* 4-3, 293-347.

Attardo, Salvatore
1997 The semantic foundations of cognitive theories of humor. *Humor: International Journal of Humor Research*, 10-4, 395–420.
Attardo, Salvatore, Christian. F. Hempelmann & Sara Di Maio
2002 Script oppositions and logical mechanisms: Modeling incongruities and their resolutions. *Humor: International Journal of Humor Research,* 15-1, 3-46.
Blake, Patricia
1975 *The Bedbug and Selected Poetry*. Bloomington: Indiana University Press.
Chabon, Michael
2007 *The Yiddish Policeman's Union*. New York: Harper Perennial.
Chandler, Raymond
1940 *Farewell, My Lovely*. New York, NY: Alfred A. Knopf.
1954 *The Long Goodbye*. New York, NY: Alfred A. Knopf.
Clark , Herbert H. & Richard J. Gerrig
1984 On the pretense theory of irony. *Journal of Experimental Psychology: General* 113(1):121-126
Cohen, Leonard
1994 *Stranger Music: Selected Poems and Songs*. New York, NY: Vintage.
Coulson, Seana
2000 *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*. New York/Cambridge: Cambridge University Press.
Dickens, Charles
1984 [1843] *A Christmas Carol*. Middlesex, UK: Puffin Books.
Fauconnier, Gilles & Mark Turner
1998 Conceptual Integration Networks. *Cognitive Science*, 22(2):133–187.
Oring, Elliott
2003 *Engaging Humor*. University of Illinois Press.
Fellbaum, Christiane (ed.)
1998 *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.
Fishlov, David
1992 Poetic and Non-Poetic Simile: Structure, Semantics, Rhetoric. *Poetics Today*, 14(1), 1-23.
Giora, Rachel
1995 On Irony and Negation. *Discourse Processes* 19:239-264.
2002 *On Our Mind: Salience, Context and Figurative Language*. Oxford: Oxford University Press.

Glucksberg, Sam
2001 *Understanding figurative language: From metaphors to idioms*. Oxford: Oxford University Press.
Grice, Herbert Paul
1978 Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics: Vol. 9. Pragmatics*. New York: Academic.
Hanks, Patrick
2004 The syntagmatics of metaphor. *International Journal of Lexicography*, 17(3):245-274.
Kumon-Nakamura, Sachi, Sam Glucksberg & Mary Brown
1995 How about another piece of pie: The allusional pretense theory of discourse irony. *Journal of Experimental Psychology: General* 124(1):3-21.
Lakoff, George
1987 *Women, fire and dangerous things*. Chicago University Press.
Lakoff, George & Mark Johnson
1980 *Metaphors We Live By*. Chicago, Illinois: University of Chicago Press.
Lloyd, John, Richard Curtis & Ben Elton
1999 *Blackadder: the whole damn dynasty*. London, UK: Penguin books.
Moon, Rosamund
2008 Conventionalized as-similes in English: A problem case. *International Journal of Corpus Linguistics* 13(1(), 3-37.
Norrick, Neal
1986 Stock Similes. *Journal of Literary Semantics* XV(1), 39-52.
Ortony, Andrew
1979 The role of similarity in similes and metaphors, in *Metaphor and Thought,* edited by A. Ortony. Cambridge, MA: Cambridge University Press.
Pollio, Howard R.
1996. Boundaries in Humor and Metaphor. Jeffrey Scott Mio & Albert N. Katz (eds.), *Metaphor: Implications and Applications*. Mahwah, New Jersey: Laurence Erlbaum Associates, 231–253.
Pope, Alexander
1813 *The Iliad of Homer*. London: J. Walker and associations.
Raskin, Victor
1985 *Semantic Mechanisms of Humor*. Dordrecht: D. Reidel.
Ritchie, Graeme
1999 Developing the Incongruity-Resolution Theory. *In proceedings of the 1999 AISB Symposium on Creative Language: Stories and Humour*, Edinburgh, Scotland.
Robbins, Tim
1990 *Skinny Legs And All*. New York: Bantam Dell.

Roncero, Carlos., John. M. Kennedy & Ron Smyth
2006 Similes on the internet have explanations. *Psychonomic Bulletin and Review*, 13(1), 74-77.
Speir, Jerry
1981 *Raymond Chandler*. New York: Frederick Ungar.
Sperber, Dan and Deirdre Wilson
1992 On verbal irony. *Lingua* 87, 53-76.
Suls, Jerry M.
1972 A Two-Stage Model for the Appreciation of Jokes and Cartoons: An information-processing analysis. In J. H. Goldstein & P. E. McGheen (eds.), *The Psychology of Humor*. New York: Academic Press, 81-100.
Taylor, Archer
1954 Proverbial Comparisons and Similes from California. *Folklore Studies* 3. Berkeley: University of California Press.
Veale, Tony
2004 Incongruity in Humor: Root-Cause or Epiphenomenon? *The International Journal of Humor* 17/4, a Festschrift for Victor Raskin.
Veale, Tony & Diarmuid O'Donoghue
2000 Computation and Blending. *Cognitive Linguistics*, 11(3-4), 253-281.
Veale, Tony, Kurt Feyaerts & Geert Brône
2006 The Cognitive Mechanisms of Adversarial Humor. *Humor: The International Journal of Humor Research*, 19-3, 305-338.
Veale, Tony & Mark T. Keane
1992 Conceptual Scaffolding: A spatially founded meaning representation for metaphor comprehension. *Computational Intelligence* 8(3), 494-519.
Whissell, Cynthia
1989 The dictionary of affect in language. In R. Plutchnik & H. Kellerman (eds.) *Emotion: Theory and research*. New York: Harcourt Brace, 113-131.
Wodehouse, Pelham Grenville (P.G.)
1955 *Bertie Wooster sees it through*. New York, NY: Simon and Schuster.
1975 *The Inimitable Jeeves*. New York, NY: Viking Penguin.