

The ABCs of XYZs: Creativity and Conservativity in Humorous Epithets

Tony Veale

Introduction

Words that evoke strong emotions or deeply-held beliefs can easily slip from the moorings of their literal meanings. Because the link between these words and the feelings they evoke is so strong, the feelings persist even when the words are used in newly minted figurative senses. Thus, while it may seem trite to call someone ‘Judas!’ for betraying a shared belief system—as Bob Dylan was memorably called at a concert in Manchester shortly after going electric—the resulting metaphor is both unambiguous and visceral. Name-calling doesn’t have to be witty or creative to be emotionally effective, and a language like English provides a range of stock names and phrases for conveying a recognisable feeling with a powerful (even if not very original) punch.

Yet originality is sometimes a necessity, especially when the feeling that is to be evoked by an epithet is neither simple nor universal—‘Judas!’ ticks both of these boxes—but complex and nuanced. Consider the comparison in (1), from a guide to wines (Skinner 2005:62):

(1) Riesling is the Kenny G of the wine world: technically brilliant, but oh-so-lacking in credibility.

This comparison is witty on a number of levels. For one, it articulates a common feeling that credibility and virtuosity do not always go hand in hand, since one can have an abundance of the latter while completely lacking the former. For another, it identifies the ideal poster-boy for this feeling, the saxophonist Kenny G, whose music has become a near-universal accompaniment to elevator rides the world over. Thirdly, it projects this identifiable feeling for a person onto an

inanimate entity for which it is surprisingly apt. Finally, it makes its point about Riesling by gratuitously inflicting collateral damage on a musician whom people love to disdain. As it happens, Kenny G is the perfect touchstone for demonstrating the validity of the complex property ‘brilliance without credibility’, and the perfect vehicle for transporting this property to a semantically-distant target.

The comparison in (1) is a figurative usage of the XYZ construction ‘*X is the Y of Z*’ (see Fauconnier and Turner 1998, 2002). But the construction also has many literal uses, as in (2):

(2) Elizabeth II is the queen of England.

Royalty is a rich vein for metaphor. For instance, the figurative XYZ in (3) is a line sung by the has-been producer Max Bialystock in the Mel Brooks musical *The Producers*:

(3) I used to be the king, the king of old Broadway.

It’s good to be the king. XYZ constructions can range from the literal, as in (2), to the mildly figurative, as in the use of ‘king’ in its conventionally metaphoric sense in (3), to the creatively figurative, as in the saxophone riff in (1).

But (3) is creative for another reason: it takes a specific instance, *Kenny G*, and converts it into a generic category of more-or-less similar notions. The conceit that the world contains many Kenny G’s in different domains of experience will seem humorous to some, and dispiriting to others. The resulting category unites a rather disparate set, which might be named ‘things which are technically brilliant but lacking in credibility’. In other words, Kenny G is used to coin a convenient label for what Barsalou (1983) calls an *ad-hoc category*. Figurative XYZ constructions excel at this two-step, allowing a speaker to both create and provocatively name a new ad-hoc category in a single bound.

In this paper we present a comprehensive analysis of figurative XYZ constructions that make incongruous use of proper-named individuals for humorous effect. We begin by considering the role of XYZs as support structures for linguistic creativity, before describing how a large corpus of proper-name XYZs is harvested from the world-wide web. We then present an analysis of the semantic and pragmatic norms that are most evident in this corpus. By establishing the conservativity of figurative XYZs in the mainstream, we can identify the most creative uses of the form among the minority of uses that violate this conservativity. The paper then concludes with some closing observations about the link between creativity and humour.

Support Structures for Linguistic Creativity

English does not provide an adjective for concisely evoking the feeling that something is technically brilliant yet lacking in credibility; the nouns ‘geek’ and ‘nerd’ come close, as do the adjectives ‘geeky’ and ‘nerdy’, but not close enough for all descriptive purposes, especially that of (1). Fortunately, popular culture perpetuates stereotypical examples of this complex property, and through a process of *catathresis* (see Black 1962), the names we give these stereotypes can be used to fill the newly identified gap in our lexicons. These catachretic metaphors often come conveniently packaged in a figurative XYZ construction, which allows a speaker to dynamically inflate a stereotypical individual into a whole new ad-hoc category of people and things. In fact, the ease with which this construction facilitates the coining of new catachretic metaphors makes the figurative XYZ a powerful support structure for linguistic creativity.

The figurative XYZ construction, and other support structures that facilitate off-the-cuff linguistic creativity, belong to a class of reusable linguistic patterns that Kay (2002) has named *patterns of coining*. These patterns allow speakers to coin well-turned descriptions that exhibit both novelty and familiarity: while the structure of each coinage is familiar— so familiar as to suggest a relatively straightforward mapping of form to meaning—its content may nonetheless be fresh and creative. For instance, the standard simile frame ‘*X is as Y as Z*’ allows a speaker to formulate a figurative comparison whose communicative goals are openly flagged (i.e., that Z exemplifies Y-ness, but X has as much Y-ness as Z) yet whose creativity depends entirely on the bindings given to Y and Z. English provides a wealth of proverbial similes with this form, as documented by (Taylor, 1954; Norrick, 1986; Fishlov, 1992; Veale and Hao, 2007 and Moon, 2008). Yet despite this abundance, speakers commonly coin their own creative similes, safe in the knowledge that should their attempt to be humorous or creative fall flat, the listener will nonetheless extract the key predication of Y-ness to X. The simile-coining pattern is not just a support-structure, but a safety-net for linguistic creativity.

Speakers sometimes subvert the simile pattern for ironic ends. When Z is an obvious counter-example of Y-ness, listeners infer that Y is just as unlikely to be a property of X. This ironic use of the simile pattern toys with the listener, establishing the expectation that Y is true of X before quickly dashing this expectation. Some ironic similes

have become conventionalized in English, such as ‘*as crazy as a fox*’ and ‘*as clear as mud*’. In each case, the dashed expectation serves as an implied criticism of, and a rejoinder to, a speaker who has suggested or implied that a situation is otherwise. For instance, the simile ‘*that lecture was a clear as mud*’ implies not just a lack of clarity, but further imparts criticism for this failure, insofar as lectures *should* be clear and understandable. Moon (2008) has observed that ironic similes are often prefixed with the hedge-marker ‘about’, while Hao and Veale (2010) argue that ‘about’ does not exclusively mark uses of irony, but is a more general marker of sardonic intent. In other words, the pattern ‘X is *about* as Y as Z’ is a support structure for sardonic creativity. Hao and Veale further note that when Y is a positive property, ‘about’ similes are almost always ironic, yet when Y is a negative property, ‘about’ similes are almost always non-ironic. In either case, ‘about’ signifies an attempt by a speaker to be humorously critical of a topic X. Whether one appreciates the irony or not, the marker subtly supports the speaker’s communicative goals, and biases the listener toward a negative and critical interpretation.

Fishlov (1992) has argued that creative comparisons are cut from much the same cloth as non-creative or clichéd comparisons. Each is dependent, in different ways, on stereotypes and their power to evoke deeply-entrenched feelings. Creative language does not abhor the use of stereotypes, just the lazy and unthinking use of stereotypes. Fishlov argues that poetic comparisons use stereotypes as building-blocks in novel combinations rather than as standalone descriptors. One might thus juxtapose or blend stereotypes to achieve humorous incongruity, e.g. to say ‘*as slow as a glacier in a traffic jam*’ rather than to reuse the cliché ‘*as slow as a glacier*’. In Fishlov’s view, non-poetic comparisons can be transformed into poetic comparisons if suitably elaborated and enriched by a speaker. Creativity is not a binary distinction: linguistic expressions can exhibit shades of creativity, and continuous elaboration and tweaking can deepen the perceived hue.

The figurative XYZ construction is a pattern of coining that allows speakers to mint their own humorous epithets. Yet few coinages that appear novel are truly original, as many XYZs simply riff upon forms that are already current. Though (1) derives its creative stretch and much of its humorous incongruity from the comparison of an inanimate wine to an inane musician, *Kenny G* has already achieved currency in the music world as a stereotype of disdained virtuosity, as shown by examples (4) – (6):

(4) Yanni is the Kenny G of keyboards. (5) Eric Clapton is the Kenny G of blues.

(6) John Mayer is the Kenny G of the guitar.

A cursory web search reveals that a wide variety of instruments—from the sitar to the banjo and the piano—each have their own *Kenny G*'s. We cannot know with certainty whether musical uses like (4), (5) or (6) logically precede (1), but it does seem likely that *Kenny G* first emerged as a stereotypical whipping boy in musical comparisons before evolving into a vehicle for describing more general, non-musical targets. In other words, the partially instantiated XYZ frame '*X is the Kenny G of Y*' has itself become a support structure for linguistic creativity, one that makes it easier to coin humorous epithets like (1). Much the same meaning is carried by each instantiation, namely that someone (or something) possesses technical virtuosity yet nonetheless fails to bridge the credibility gap. This particular XYZ allows a respondent to quickly coin a context-specific riposte to the argument that virtuosity is in itself a sufficient cause for acclaim.

If the precise import of Y to X is unclear, an explanation can always be appended, as in (1), without diminishing the humour of the comparison. Whereas jokes typically suffer for having to be explained, the humour carried by constructions like the simile frame and the figurative XYZ is often heightened if profiled by a concise explanation. In these cases the figurative description acts more like a humorous riddle than a narrative joke, since it is the explanation that delivers the humorous punch. Studies of simile usage on the web—where users often coin their own throwaway comparisons—show that most novel web similes are accompanied by an explanation that drives home the point of the comparison (see Roncero, Kennedy and Smyth 2006). The availability of an explanation means that speakers can be more daring and creative in their comparisons, safe in the knowledge that the right meaning will be conveyed even if the humorous conceit should fall flat.

A well-crafted XYZ can offer the listener a multimedia experience, evoking not just attitude, but imagery, sound and even taste. The comparison in (1) thus complements our taste memories of Riesling with the musical strains of *Kenny G*. The point, of course, is to make an abstract property seem tangible to our senses. The comparison in (7) below evokes the artificial taste of an ersatz soda to make its point:

(7) He [*Edward Cullen*] is the Diet Coke of vampires.

We see here one portrayer of vampires—the actor Stephen Moyer

of the TV series *True Blood* – disparage a vampire played by another actor in another drama (The *Twilight* saga’s Edward Cullen). The claim is that Cullen is as much an ersatz vampire as Diet Coke is an ersatz cola, or to reformulate the metaphor, that Cullen is merely vampire-*lite*. The construction ‘the Diet Coke of evil’ had earlier been popularised in the film *Austin Powers: The Spy Who Shagged Me*, and the partially instantiated construction ‘*X is the Diet Coke of Y*’ has since become a widely-used support structure for diminishing X’s credentials as an example of Y-ness. Just as (1) evokes the musical strains of Kenny G, (7) and its ilk evoke the chemical taste of sugar substitute to make their meanings all the more palpable to the listener.

Proper-names like *Kenny G* and *Diet Coke* have become brand-names for assorted commercial offerings, but also serve as instantly recognisable place-holders for specific attitudes, beliefs and sentiments. Since proper-names are easy to identify via their capitalisation patterns, we focus here on the harvesting and analysis of figurative XYZ constructions where the Y element is a proper-named individual. We begin by describing the web harvesting process in the next section, before discussing our findings.

Harvesting A Large Web Corpus of XYZs

Figurative uses of the XYZ construction carry no special marker to indicate their metaphoricity. However, if Y is a proper-named individual—and this is easily ascertained on orthographic grounds—then it is incongruous to use Y to denote a category of individuals. Incongruity theories of humour suggest that this category error must be resolved by showing the incongruity to be appropriate on some higher level, so that Y can be understood as an appropriate designation for a whole category of Y-like entities. Figurative XYZs of this kind are obvious examples of Glucksberg’s category inclusion theory of metaphor (see Glucksberg 1998, 2001), since Y does not literally denote itself in the XYZ, but stands in place of a category in which Y is a salient member and to which X can likewise be added.

A broad sample of proper-name XYZ coinages can be acquired from the web, though as Veale and Hao (2007) note in the context of harvesting web similes, the process is not without complexity. Popular search engines like Google offer the broadest and most up- to-date coverage, but their query languages lack the finesse of dedicated corpus- processing tools. Moreover, commercial search engines may

rank their results by popularity, or authoritativeness, or even by their perceived relevance to a paying advertiser, but not by diversity: a simple query like ‘* is the * of *’ will not only match a great many phrases that are not figurative XYZs, but will also retrieve a great many duplicates, overlooking interesting one-offs that are used in obscure blogs or little-visited web-sites. The remedy, as Veale notes, is to dispatch an enfilade of specific, partially-instantiated queries rather than a single, over-arching query. Following our discussions in the previous sections, we should thus identify the most useful Ys for coining XYZs, and formulate queries for the corresponding partially-instantiated support structures, such as ‘* is the Kenny G of *’.

The Google database of web ngrams (Brants and Franz 2006) is a large collection of the most common English word sequences on the world-wide web. Each ngram contains between one and five words, is case-specific (allowing for orthographic matching), and has a guaranteed web frequency of 40 or more documents. Looking to the Google 4-grams, we can thus identify all matches for the pattern ‘the *Fname Lname* of’, where *Fname* and *Lname* match any capitalised words that might plausibly be used as the first and last parts of a name respectively. A large list of allowable name elements is harvested from Wikipedia for this task. Likewise, a collection of evocative one-word names (such as *Mozart*, *Einstein* and *Napoleon*) is used to find matches in the Google 3-grams. The combined matches for 3-grams and 4-grams provide instantiations for the proper-named Y component of each XYZ query, allowing us to automatically dispatch a corresponding Google query for each specific Y. In return, Google provides a set of up to 200 matching text snippets for each query, such as the snippet in (8), which contains an extract from the *Sunday Times* that is retrieved in response to the query ‘is the Picasso of’.

(8) 28 Mar 2010 ... ‘Ferran is the Picasso of the modern kitchen,’ enthuses Rafael

Anson, president of the International Academy of Gastronomy. An automated filter is then applied to each snippet to identify those that contain well-formed X, Y and Z components, and to exclude those whose X is a pronoun like ‘he’ or ‘it’, as well as those whose Z component does not conform to any of a set of simple patterns (such as a bare noun, like ‘crime’ or ‘wines’, a temporal or geographic specifier, like ‘European Union’ or ‘20th Century’, or a phrase of the form ‘the Z’, ‘the Z world’, ‘the Z domain’, ‘the Z genre’ or ‘the Z industry’). In all,

a series of over 3000 queries harvests more than 60,000 snippets from Google, but subsequent filtering pares these down considerably, to extract a corpus of 2190 unique XYZs for 668 different Ys.

As expected, the most frequent Ys are all prominent individuals whose propensities and abilities are well established in popular culture. The 20 most frequent Ys account for about 10% of all XYZs in the corpus, and the individual-types in this top 20 are broadly representative of the whole, ranging from historical figures to fictional figures to artists, musicians, actors, politicians and sportsmen. These individuals are ciphers for some commonly ascribed properties in humorous epithets: *Benedict Arnold*, for instance, stands for any individual who is traitorous or just plain fickle about which side to support; *Rush Limbaugh* can stand for any political loud-mouth with partisan views; and *Chuck Norris*, an expressionless actor who is lampooned relentlessly on the internet, can stand for any person or thing that is equally uncomplicated and uncompromising in its behaviour.

With 21 examples, *Michael Jordan* is the most commonly used Y in our corpus. As an athlete who occupies the pinnacle of his chosen sport, Jordan has become a role model for strivers in any sport, and some non-sports besides. Here are the 21 Xs that are compared to Jordan in our corpus (with the corresponding Zs in parentheses):

Manny Pacquiao (Philippines), Andrew Gaze (NBL), Chet Snouffer (boomerang), Garry Kasparov (chess), Mwadi Mabika (WNBA), Vince Young (NFL), Pádraig Harrington (golf), Tiger Woods (golf), Randy Couture (martial arts), Daryll Pomey (Philippines), Tony Hawk (skateboarding), Champ Hallett (wheelchair basketball), David Berg (courtroom), Bronwyn Weber (cakes), Michael Chabon (literary), the tuna sandwich (mid-day meal), Billy Bob Thornton (movies), Ralph Appelbaum (museums), Allan Bloom (seminars), Britney Spears (pop), Randall Ross (rare books)

Some Xs are closer to Y than others, but most are themselves sportsmen. The closest are basketball players from other leagues (the NBL) or of another gender (Mwadi Mabika of the WNBA). Yet Jordan is a model of excellence for any competitive endeavour, and has leadership qualities that even tuna lovers can apparently find inspirational.

The corpus contains 1312 different Zs, once support words like ‘world’, ‘genre’, ‘domain’ and ‘industry’ have been stripped out, though the top 20 most frequent Zs account for 367 different XYZs, or more than 16% of the whole. Once again we find the same conceptual diversity in

this top tier, which includes times (the 21st and 20th centuries proving most popular), geographic locations (North, South, East and West are all in the top 20 Zs), sports and sports organisations (the NBA, NFL and NHL), politics and political parties (the Republicans, the GOP, the Democratic Party) and political orientations (Left versus Right), as well as art, music and gaming. The most frequent Z, 21st Century, accounts for 66 different XYZs in the corpus.

We must supplement this superficial breakdown of XYZs with an additional layer of semantic generalization if we are to identify the semantic norms that dominate the corpus or quantify the extent to which they hold sway. These generalisations, discussed in the next section, will allow us to explore the balance of conservativity to creative incongruity in our corpus of figurative XYZs.

Drilling Down: Conservativity in Figurative XYZs

Any system of semantic generalisations should be suited to the corpus at hand. Consider the following sampling of figurative XYZs from the Web, as found in our Web corpus:

- (9) (a) *Paris Hilton is the Zsa Zsa Gabor of the 21st Century* (b) *Victoria Williams is the Yoko Ono of the folk scene* (c) *Chris Manion is the Woody Guthrie of the right*
- (d) *Qifa Nabki is the Winston Churchill of the Islamic Resistance*
- (e) *Nick Denton is the William Randolph Hearst of the blog world*
- (f) *Pdq Bach is the Weird Al Yankovich of the classical music world*
- (g) *David Wetherell is the Warren Buffet of the internet*
- (h) *Steve Jobs is the Walt Disney of the tech world*
- (i) *Ben Bernanke is the Tony Robbins of the financial world*
- (j) *Newt Gingrich is the Trotsky of the Hard Right*
- (k) *Roger L'Estrange is the Torquemada of the late Stuart age*
- (l) *David Cameron is the Tony Blair of the conservative party* (m) *Michael Jordan is the Tony Hawk of the basketball world* (n) *Milton Caniff is the Rembrandt of the comics*
- (o) *Scipio Africanus is the Tommy Franks of the Roman legions*
- (p) *Peter Brett is the Tolstoy of the F train*
- (q) *Daniel Melingo is the Tom Waits of the contemporary tango*
- (r) *Shahrukh Khan is the Tom Cruise of the Bollywood Industry*

- (s) *Edward Abbey is the Thoreau of the desert*
 (t) *June Wanniski is the Thomas Paine of the Reagan revolution*
 (u) *Bill Gates is the Thomas Edison of the tech industry*
 (v) *Nicholas Sparks is the Stephen King of the mush-brained romantic novel*

Here we see musicians compared to musicians, writers to writers, artists to artists, athletes to athletes and businessmen to businessmen. Overall, even in the most humorous cases, one can sense a tremendous conservativity in this sample, where Xs and Ys seem to be drawn from a relatively small inventory of stock semantic types or domains. To quantify the degree to which this intuition applies to the corpus as a whole, we need to move from the level of tokens to the level of types, and semantically annotate each XYZ.

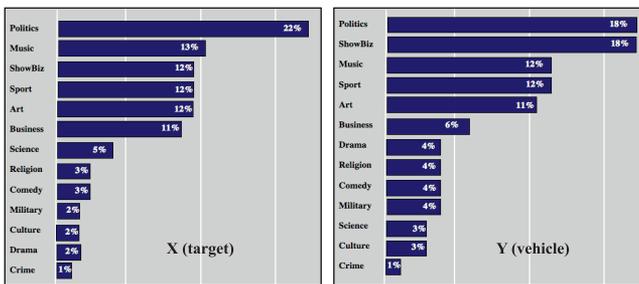


Figure 6.1. Relative distribution of domain annotations for Xs (left) and Ys (right).

More specifically, we need to annotate each X and Y element with one or more semantic types that are indicative of the domains from which they have been plucked. Such a scheme should, *a priori*, seem sensible and well-motivated, while any noteworthy observations arising from the resulting analysis should, *a posteriori*, not be attributable to an artefact of the tagging scheme. An exploratory pass through the data leads us to choose a system of 13 distinct types for annotating the Xs and Ys in XYZs: *Politics, Music, Art, ShowBiz, Military, Crime, Business, Religion, Sport, Comedy, Culture, Drama, Science*. This set does a good job of capturing the diversity of the corpus with an acceptable level of generality. *Business*, for instance, covers the worlds of commerce, finance and industry, while *ShowBiz* covers real individuals who perform on TV, on stage, or in movies. In contrast, *Drama* is used to annotate fictitious characters who appear in movies, books or other

narrative forms. *Culture* is used to annotate individuals that represent different ethnic or social groups, as well as historical figures that contribute to our understanding of a particular culture.

The relative distribution of these semantic annotations for Xs (left) and for Ys (right) is shown in Figure 6. 1. *Politics* is the most popular annotation for both components of our XYZs (22% as a target of description in the X position vs. 18% as a vehicle of description in the Y position). The distribution of other annotations shows just slight variation between Xs and Ys.

It is worth looking at the most popular annotation, *Politics*, in greater depth. Figure 6. 2 shows the relative distribution of annotations for X when the Y component of an XYZ is annotated as belonging to the *Politics* domain.

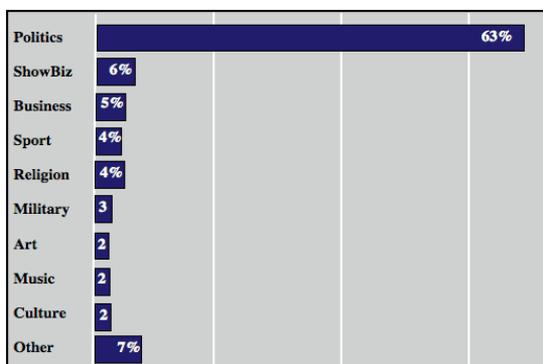


Figure 6.2. Relative distribution of annotations for Xs when Y is in the *Politics* domain.

Politics appears to be an incestuous domain, conceptually at least, and we can see from Figure 6. 2 that political vehicles, like *Trotsky* in 9(j) and *Tony Blair* in 9(l), are predominantly used to describe political targets, like *Newt Gingrich* and *David Cameron*. One can ask whether this apparent domain conservativity – wherein Xs and Ys tend to belong to same semantic domains – is a feature of the corpus as a whole, or whether *Politics* is a special case. The matrix in Figure 6. 3 provides a breakdown of all domain-to-domain mappings in our corpus of figurative XYZs. Each row in Figure 6. 3 corresponds to a different semantic domain as used to annotate a Y component, while each column presents the distribution of semantic annotations

for the X components it is used to describe. Note, for instance, that the intersection of the row *Business* and the column *Sport* contains the number 15: this indicates that, in 15 of the XYZs where the Y component is annotated with *Business*, the X component is annotated with *Sport*. In other words, 15 of the *Business* metaphors in our corpus are *Sport-is-a-Business* metaphors.

	<i>Politics</i>	<i>Music</i>	<i>Art</i>	<i>ShowBiz</i>	<i>Business</i>	<i>Crime</i>	<i>Military</i>	<i>Sport</i>	<i>Comedy</i>	<i>Religion</i>	<i>Science</i>	<i>Culture</i>	<i>Drama</i>
<i>Politics</i>	296	10	14	30	25	1	15	22	4	20	10	10	4
<i>Music</i>	22	196	19	24	14	0	1	21	3	5	6	0	4
<i>Art</i>	10	24	177	15	17	0	0	14	3	4	5	1	2
<i>ShowBiz</i>	43	40	37	170	42	2	0	34	27	9	10	9	14
<i>Business</i>	14	5	9	9	87	0	0	15	0	0	14	2	4
<i>Crime</i>	6	1	2	0	5	9	0	3	1	2	4	2	0
<i>Military</i>	46	1	9	2	10	0	32	3	1	5	4	0	0
<i>Sport</i>	15	20	3	24	22	1	2	153	9	4	7	1	3
<i>Comedy</i>	18	5	5	8	7	1	0	12	28	3	1	1	3
<i>Religion</i>	34	6	2	4	8	2	3	4	1	26	4	5	0
<i>Science</i>	6	1	5	2	12	2	1	4	1	4	46	5	1
<i>Culture</i>	13	6	3	10	7	0	1	0	1	1	2	18	2
<i>Drama</i>	19	9	6	10	25	2	2	15	2	3	2	8	21

Figure 6.3. Mappings of Y domains (rows) to X domains (columns). Each row is a different semantic domain for a Y, each column a different semantic domain for an X, where each cell contains the number of XYZs with this combination of X and Y domains.

The largest value in each row in Figure 6.3 is highlighted with a dark circle. With the exception of *Politics-as-Military* (46 *Military* metaphors), *Politics-as-Religion* (34 *Religious* metaphors) and *Business-as-Drama* (25 *Drama* metaphors), note how these dark highlights mainly occur on the diagonal, indicating that domain conservatism is a strong convention of figurative XYZs.

It turns out that web XYZs are also deeply conservative in the more traditional reading of the word ‘conservative’, for they are remarkably male-centric too, with very few female concepts on either side of the equation. Figure 4 shows the breakdown of XYZs by gender on an X-to-Y basis.

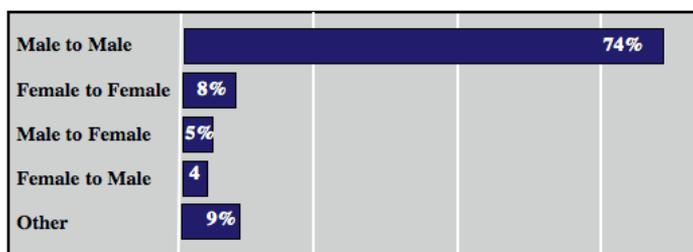


Figure 6.4. Breakdown of source-to-target mappings by gender of the mapped entities.

Other denotes expressions where either the source or the target has no obvious gender.

It certainly is a man's world, for XYZs at least, and though just 9% of our corpus involves cross-gender XYZs, there are more females in these comparisons than there are in the pure female-to-female cases. Figurative comparisons can bridge large semantic gaps between domains, but gender appears to be a bridge too far for most comparisons.

Most of our XYZs involve individuals from the real world, and so very few exploit fictional characters in either a source or target capacity. However, in XYZs that do draw upon the world of fiction, fictional sources are four times more likely than fictional targets, and as shown in Figure 6.5, a small percentage of XYZ expressions employ fictional sources *and* fictional targets.

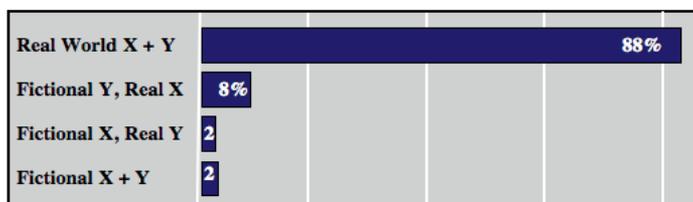


Figure 6.5. Breakdown of source and target entities by fictional status.

You might think it impossible to hurt the feelings of a fictional character, but the web XYZ 'Bruce Wayne is the Donald Trump of the DC universe' comes close. One is certainly as rich as the other, though Batman might have trouble squeezing Trump's meringue-shaped hair into his cowl. A web XYZ that describes a real person in terms of a

fictional entity is ‘*Jann Wenner is the Charles Foster Kane of the baby boomers*’. Wenner, the founder and publisher of *Rolling Stone* magazine, might just as easily be compared to William Randolph Hearst, the real-world newspaper magnate on which *Citizen Kane* was based. More flatteringly, our corpus reveals that Warren Buffet is considered by some to be ‘*the Sherlock Holmes of the stock market*’, while right-wing FOX news host Glen Beck has been called ‘*the Homer Simpson of the airwaves*’. Since the same heroic archetypes continually resurface in popular culture, one fictional character will sometimes be compared to another, as in ‘*Allan Quatermain is the Indiana Jones of the Victorian age*’ and ‘*Jack Sparrow is the Han Solo of the Caribbean*’.

Speakers also tend toward conservativity when it comes to time. For instance, we often compare a contemporary individual to a historical entity, as in ‘*Rupert Murdoch is the William Randolph Hearst of the 21st Century*’. But one can also compare a historical entity to a contemporary individual, as in ‘*Jefferson is the Trotsky of the 18th Century*’ or ‘*Russ Meyer is the Tarantino of the 70’s*’. Overall, there is a very strong preference in XYZs for the backward-looking comparison, in which a contemporary individual is compared to a similar individual from the past. As shown in Figure 6, future-looking comparisons (such as ‘*Lillie Langtry is the Lindsay Lohan of the late 19th century*’) account for less than one third of all *TimeToTime* mappings in our corpus of web XYZs.

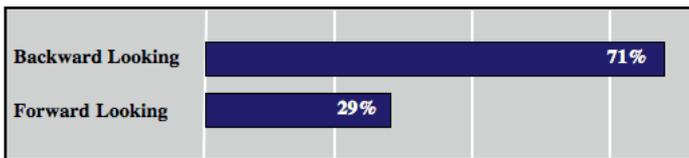


Figure 6.6 Breakdown of *TimeToTime* comparisons according to whether they are backward-looking (*X* lives after *Y*) or forward-looking (*Y* lives after *X*).

It seems that even when we strive to be creative we still cling to established conventions when it comes to domain, gender and time.

Creativity in Figurative XYZs

Let’s conclude our discussion of XYZs by looking at the examples that do not yield so neatly to our analysis thus far. Some of the most humorous XYZs in our corpus exploit a source individual *Y* to describe a non-human object, as in the following cases in (10):

- (10) (a) *Chico's Tacos is the Willy Wonka of culinary experiences*
 (b) *Der Sturmer is the Rush Limbaugh of the Third Reich* (c) *The Woodland is the Rodney Dangerfield of the lineup* (d) *DSL is the Rocky Balboa of the fast-access future*
 (e) *File Organizer is the Rocky Balboa of the genre*
 (f) *Alfa Romeo is the Quentin Tarantino of the automotive world*
 (g) *Facebook is the Patrick Henry of the 21st Century*
 (h) *Apple's iThingy is the Paris Hilton of mobile phones*
 (i) *Chicken Inasal is the Oprah Winfrey of the menu*
 (j) *Nintendo is the Ned Flanders of the console world*
 (k) *The tuna sandwich is the Michael Jordan of the mid-day meal*
 (l) *Samsung DLPs are the Lindsay Lohan of the television market*
 (m) *Pac Man is the King Lear of the 1980's 8-bit videogame revolution*
 (n) *The Razr is the Kate Moss of phones*
 (o) *The Borgata is the Julia Roberts of casinos*
 (p) *Bradley's Battleship is the John Travolta of board games*
 (q) *Krug is the Dorian Gray of the wine world*
 (r) *Red meat is the Donald Trump of cancer*
 (s) *Copy-protected CDs are the Dick Cheney of the music industry*
 (t) *Tungsten is the Cleopatra of the elements*
 (u) *Platinum Pro is the Chuck Norris of the editing world*
 (v) *The K750i is the Chuck Norris of the photography world*
 (w) *Toyota Prius is the Che Guevara of the [eco-friendly car] movement*
 (x) *The Manhattan is the Cary Grant of cocktails*
 (y) *Big Bordeaux is the Barry Bonds of the wine world*
 (z) *The Montrachet is the Angelina Jolie of the pack [wines]*

We also find the following XYZs in (11) that use people to describe an animal or a plant:

- (11) (a) *The Parrot is the Robert De Niro of the bird world*
 (b) *Moby Dick is the Samson of the ocean*
 (c) *The Blue Marlin is the Muhammad Ali of the fish world*

- (d) *Pit Bulls are the Mike Tyson of the K9 world*
- (e) *The Boxer is the George Clooney of the dog world*
- (f) *The Northern Pintail is the Audrey Hepburn of the duck world*
- (g) *Whomping Willow is the Mike Tyson of the plant kingdom*
- (h) *The potato is the Tom Hanks of the vegetable world*

These figurative XYZs resemble similes more than analogies, since most are built around a single highly-salient property of the source concept. So, for instance, Tom Hanks is *versatile*, Rocky Balboa is *resilient*, Chuck Norris is *implacable*, Muhammad Ali is *graceful*, Kate Moss is *super-slim*, Donald Trump is *aggressive*, Angelina Jolie is *voluptuous*, Audrey Hepburn is *elegant*, Cary Grant is *sophisticated* and Rodney Dangerfield *gets no respect!* As with humorous similes, a comparison can seem flimsy and gratuitous if it aligns two very different concepts from distant parts of our conceptual systems on the basis of a single shared property, especially if this property has different meanings in the source and target domains. Such flimsiness makes it hard to take a comparison seriously, but can make it easier to recognise the humorous intent behind the comparison. As a result, we still comprehend the core message—that X has the salient property stereotypically associated with Y— while the incongruous juxtaposition gives us the added bonus of a smile. Humorous XYZs are double-edged comparisons that can cut both ways, since information inevitably flows in both directions, from Y to X (the real message) *and* from X to Y (the humorous bonus), to meet in the middle to construct a blended mental image. For instance, when we imbue the humble potato with the versatility of the actor Tom Hanks, we might imagine the different culinary uses of potatoes as the different roles that a talented potato can play in a meal. In doing so, we effectively equate Hanks with a talented potato, slyly diminishing the value of his craft.

Even flimsy XYZs have the potential to be elaborated into more complex analogies. Barry Bonds is famous as a *'big hitter'* in baseball, so any wine compared to this high-profile player might also be said to *'pack a wallop'*. But Bonds has also been dogged by accusations of doping, and these allegations can transfer to the wine domain as suspicions of chemical adulteration. Consider the comparison *'Red meat is the Donald Trump of cancer'*, which reads like a puzzle that does not want to reveal its meaning too easily. We know, for instance, that Donald Trump is a famously aggressive property tycoon, and that cancer is scariest when it spreads aggressively. We also know that

aggressive predators tend to be voracious consumers of red meat, but these two puzzle pieces refuse to click together. Something is missing, and the meaning only becomes clear when we look to the explanatory text in which the original author imbeds the XYZ. Vegan (2007) notes that since red meat has been implicated in the development of many different kinds of cancer it can be metaphorically categorised as an aggressive and opportunistic builder of cancers. So 'the Donald', who is the very model of an aggressive and opportunistic property developer, perfectly fills the Y role in this XYZ. Humorous XYZs provide attention-grabbing mental imagery but do not always wear their meanings on their sleeves. Yet their humour is scarcely diminished by having to explain these meanings; rather, the imagery and the explanation play complementary roles in a text, with the former serving as a sturdy and memorable scaffolding for the latter.

Concluding Remarks

In a rather nice retelling, the *Time Out Guide to Mumbai & Goa* (Mirani 2008) presents the tale of how the Persian Parsis came to settle in India. Here is a brief extract:

They [the Parsis] arrived in Gujarat in the eight or ninth century and sought asylum from the local king. He is said to have sent them away with a glass of milk full to the brim – his way of saying that his kingdom was full. The Parsi elders conferred, added some sugar to the milk and sent it back—to suggest that they would mix thoroughly and sweeten the life of the community.

The Parsi elders were wise enough to recognise a creative symbol when they saw one. Rather than view the milk as just a beverage, they perceived instead its new figurative meaning. The concoction is not just a physical blend (of milk and sugar), but a figurative blend too (of Gujarati and Parsi). It also implies a simile (the Parsis promise to be 'as sweet as sugar'), a metaphor (the glass suggests a *state as container* conceptual metaphor) and an analogy (the Parsis will be to Gujarat 'as sugar is to milk'). Even everyday ingredients can yield a creative wallop when they are used in the right way.

This tale of the Parsis is a story of creative *non-linguistic* communication, but its lessons are just as applicable to verbal communication: to be creative with language is to make the most of what we are given, to recognize and unlock the full potential of the commonplace and the familiar. In this paper we have explored how a

creative speaker can extract new value from old tropes, and how can one can exploit the most familiar stereotypes in compositions that are as meaningful as they are surprising. We have seen how speakers can turn the most prominent individuals in a culture, whether real or fictional, historical or contemporary, into symbols of something bigger than themselves.

Most surprising of all, these speakers most often achieve their creative ends while having one foot squarely and conservatively planted in the realm of the normative and the conventional. To conclude with an XYZ from the economics domain, creative speakers are the *investors* of the language world: they derive a return on investment from what is plainly accessible to everyone, yet which is all too often hidden from plain sight.

References

- Barsalou, Len. 'Ad-hoc categories,' *Memory and Cognition* 11(3): 211–27, 1983.
- Black, Max. *Models and Metaphors: Studies in language and philosophy*, Ithaca: Cornell University Press, 1962.
- Brants, Thorsten and Franz, Alex. *Web 1T 5-gram Version 1*, Linguistic Data Consortium, 2006.
- Fauconnier, Gilles and Turner, Mark. 'Conceptual integration networks,' *Cognitive Science*, 22(2):133–87, 1998.
- . *The Way We Think: Conceptual blending and the mind's hidden complexities*. Basic Books, 2002.
- Fishlov, David. 'Poetic and non-poetic simile: structure, semantics, rhetoric,' *Poetics Today* 14(1): 1–23, 1992.
- Glucksberg, Sam. 'Understanding metaphors,' *Current Directions in Psychological Science* 7: 39–43, 1998.
- . *Understanding Figurative Language: From metaphors to idioms*, New York: Oxford University Press, 2001.
- Hao, Yanfen and Veale, Tony. 'An ironic fist in a velvet glove: creative misrepresentation in construction of ironic similes,' *Minds and Machines* 20(4): 483–88, 2010.
- Kay, Paul. *Patterns of Coining*. Unpublished manuscript, 2002.
- Mirani, Leo. *Time Out Guide Mumbai & Goa*, Newport: Ebury Press, 2008.
- Moon, Rosamund. 'Conventionalized as-similes in English: a problem case,' *International Journal of Corpus Linguistics* 13(1): 3–37, 2008.
- Norricks, Neal R. 'Stock similes,' *Journal of Literary Semantics* XV(1): 39–52, 1986.

- Roncero, Carlos, Kennedy, John M. and Smyth, Ron. 'Similes on the internet have explanations' *Psychonomic Bulletin and Review* 13(1): 74–77, 2006.
- Skinner, Matt. *Thirsty Work* London: Hardie Grant Books, 2005.
- Taylor, Archer. 'Proverbial comparisons and similes from California' *Folklore Studies* 3, Berkeley, CA: University of California Press, 1954.
- Veale, Tony and Hao, Yanfen.. 'Comprehending and generating apt metaphors: a Web-driven, case-based approach to figurative language', *Proceedings of AAAI-2007, the 22nd AAAI conference on Artificial Intelligence*, 2007.
- Vegan, Eccentric (blogger pseudonym). 'Red meat is the Donald Trump of cancer. <http://www.vegansoapbox.com/red-meat-is-the-donald-trump-of-cancer/>, 2007.