Using Type to Add Data to Data Visualizations

Richard Brath and Ebad Banissi

Abstract—Type attributes, such as bold and italic, can be used to represent data in visualizations. A review across domains shows various uses of type attributes and these can be itemized and assessed for different ways to encode data. Numerous examples show how these attributes can be applied to new kinds of visualizations.

Index Terms—Font attributes, Text visualization, Alphanumeric glyphs, Quantitative typography

INTRODUCTION

Data visualization is the process of turning data into visual attributes, such as the size of a bar on a bar chart, the color of a county on an election map, or the location of a dot on a scatterplot. Unfortunately, most visualization designers today do not leverage typography to express data using type attributes such as bold, italic, condensed and so on. Visualization textbooks identify visual attributes such as size, hue, intensity and texture but relegate text to a single entry (e.g. [1,2,3,4,5]). Type is an afterthought and many visualizations are typographically mute.

For example, figure 1 shows two highly popular data visualizations. Top is a treemap visualization: this treemap is a visualization of the stock market where each box represents a company with the size and color indicating the company’s stock market size and the change in its stock price. Boxes are organized by sector with simple labels added on top.

The lower image in fig. 1 is a tag cloud. Type is important in this visualization - the size of the word is related to the frequency of the word in the original text. However font color and typeface are arbitrary and are not associated with any data.

Instead, the design space of visualization could be much richer harnessing typography to express data encoded directly into text. There are many possibilities that can result from design exploration.

There are three objectives for this research:
1. Explore type history for examples where differences in typographic attributes convey data.
2. Catalogue those attributes and how they could be used.
3. Invent new kinds of typographic visualizations.

1 EXPLORE BACKGROUND

There are many domains that have historically used type to express data. The most obvious may be cartography (e.g. [8,9]), for example, road maps that use all caps to indicate state capitals. The map in figure 2 uses typographic attributes such as reverse italics to indicate rivers, underline style to indicate different administrative levels of cities, as well as all caps, spacing and typeface.

Fig. 2. Map with font size, caps, spacing, various underline styles, and italics/ reverse italics indicating data [10].

In typography, there are many examples. The family tree in figure 3 uses bold to indicate major branches, all caps to indicate regions, small caps to represent rulers, roman text for descendants and italics for spouses.

Fig. 3. Genealogical diagram from 1820 with bold, all caps, small caps, roman and italics indicating data [11].

The typographic table of contents in Chambers’ Cyclopaedia from 1728 uses italics for major branches of science; small caps for specific topics, and superscripts for a chapter number (figure 4).
Haeckl’s diagram of the Pedigree of Mammals uses various typefaces to differentiate between common names, scientific names and so forth for various branches of mammals.

Other domains also use typographic attributes to encode data within text [see 14], including:

- **Notation Systems** such as mathematical formulas (e.g. \( \mu(A) = \inf\{\lambda(O) \mid O \in \Theta, A \subset O\} \)), chemical formulas (e.g. \([\text{As}(\text{Ni})\text{Zn}]^2\)), and markup notation (e.g. `<div class="body">Text</div>`)
- **Alphabetic Charts** for different domains (e.g. [12,17,18], including:
  - **Typefaces** for different domains
  - **Data Visualization** sometimes does use text with typographic attributes. E.g. Skupin visualizes knowledge domains using cartographic labelling techniques such as font size, orientation and spacing [15].
  - **User Interface Design** historically recommended against using font attributes (because of low resolution displays) although current web design references now often recommend using font attributes and modern code editors use type attributes such as bold, underline and italics to highlight programming syntax.

### 2 Catalogue of Type Attributes

Based on an analysis of type use across domains and typography texts (e.g. [16,17,18,19]), potential font attributes can be identified as outlined in table 1. The second column lists the font attributes with the first column a grouping. The final attributes at the bottom indicate probability of preattention [21].

<table>
<thead>
<tr>
<th>Group</th>
<th>Font Attribute</th>
<th>Best (for encoding):</th>
<th>Example</th>
</tr>
</thead>
</table>
| Glyphs | Alpha-numeric Text Glyph | O, G | G
| | Symbols | G | using " emphasizes |
| | Oblique / Italic | G | using "slopes angle |
| | Case (no small caps) | G | possible O (O-3 levels) |
| | Typeface | G | G, Q (using lengths) |
| | Condensed | G, Q | G, Q (using lengths) |
| | Squished | G, Q | G | |
| Sequence | Spacing | G, Q | G |
| | Super/subscript | G, Q | G |
| | Delimiters | G, Q | G |
| Font Design? | X-height | G, Q | G |
| | Contrast / Stress angle | G, Q | G |
| | Serif length / Bracket size | G, Q | G |

Table 1. Font attributes and types of encoding.

The third column indicates how these attributes can encode data as identified by information visualization researchers and cartographers. For example, quantities, can be encoded using font weight or obliques; while typeface is best used to encode different categories. The final column shows some examples of each attribute.

Some visual attributes visually pop-out from their surrounding, being perceived almost immediately regardless of the number of other items (figure 6) [20].

In psychology, these visual channels are called pre-attentive and some of these attributes have a stronger pop-out effect than others [21] (in typography, this effect is referred to as color, where the objective is to reduce any pop-out effect in a well-designed font). In turn, each font attribute can be related back to these preattentive attributes, to get a sense of which may be more effective at visually standing out (table 2). For example, font weight, using visual channel of intensity and size is likely to be more effective at visually standing out than case. Although size is preattentive, serif size or bracket size may be too small to be noticeable, and further small changes may be subject to change blindness.

3 New Visualizations

Given a set of font attributes, the next question is how they might be used to create new kinds of visualizations.

### 3.1 Skimming Text

Text skimming is a reading technique of rapid eye movement across a large body of text to get the main ideas and content overview. At a low level, the strategy requires the reader to dip into the text looking for proper nouns, unusual words, enumerations, etc. To make uncommon words pop-out, word frequency analysis can be used to identify the least common words and then make these have the heaviest weight. Figure 7 shows the opening paragraph of *The Wizard of Oz* formatted for skimming.
Italics are also used for less important parts of speech (e.g. articles, pronouns) to create greater differentiation from the heaviest weight text and enhance figure-ground separation.

Dorothy lived in the midst of the great Kansas prairies, with Uncle Henry, who was a farmer, and Aunt Em, who was the farmer’s wife. Their house was small, for the lumber to build it had to be carried by wagon many miles. There were four walls, a floor and a roof, which made one room; and this room contained a rusty looking cookbook, a cupboard for the dishes, a table, three or four chairs, and the beds. Uncle Henry and Aunt Em had a big bed in one corner, and Dorothy a little bed in another corner. There was no garret at all, and no cellar except a small hole dug in the ground, called a cyclone cellar, where the family could go in case one of those great whirlwinds arose, mighty enough to crush any building in its path. It was reached by a trap door in the middle of the floor, from which a ladder led down into the small, dark hole.

Fig. 7. First paragraph from The Wizard of Oz formatted for skimming.

3.2 Reviewing Opinions

The movie review website, Rotten Tomatoes, has potentially hundreds of reviews per movie. Attempting to get an overall sense of the reviews across some different movies is difficult.

Instead, figure 8 shows a list of movie reviewer’s quotes for two movies. The amount of bold along each line indicates the overall reviewer’s score - a short amount of bold is a poor score while full bold indicates the highest score. For a single movie, one can see what the range of reviews are - e.g. how strong is the difference of opinion between the best and worst reviews; and also facilitates comparison across reviews - for example a review that has more total bold is a higher scoring movie overall.

Despicable Me 2

This is a sequel that’s even less necessary than Monsters University; often times it fits. Despicable Me, the animated supervillain comedy from 2010, was an average flick we

Given the outlandish premise, you’d wish the film twinkled with a more savory sense. For cynics and detractors, it may at time feel like this sequel exists for nothing more than

Cute family fun, but lacks the pop of the original. Gru has gone from despicable to dr

Its hyperactive vibrancy is universally boredom-good.

Not a great movie for sure, but if your kids want to see this there is enough humor to

Gru still has charm and kids will adore the Minions.

Steve Carell’s Slavic inflections as Gru in the trick, as before. Wigg’s clever translator

Once again, there’s nothing here that’s particularly original or memorable, but the ch

The film easily surpasses the original, while leaving room for further sequels.

An animated sequel that, despite not achieving the inspired legacy of the first movie

Parts James Bond flick, ‘Get Smart’ episode and Pixar-esque family adventure, ‘Des

Ranks as one of the best animated sequels of all time.

Though jammed-up with too much pointless plot, Despicable Me 2 remains one of t

Not as consistently funny as the original, Despicable Me 2 still proves it’s self a quencher

The prequel, gizmos and Looney Tunes ‘Violence’ will elicp giggles from kids while it

How to Train Your Dragon

Everything from the angle of the shot to the speed of the editing projects an end view

Here, Viking life is grim, hostile and heavy with social pressure - kind of like Gossip

The visuals are striking, the script sharp and well-paced and it all wraps up with a bit

Full of wonder, charm and dragons not doubling as stand-up comics.

It’s a brisk, amusing piece that doesn’t have the weary ween that leaves a tinge.

Following a slow, overly verbal start, this dragon tale takes off!

Baruchel, Ferguson and Butler supply a contagious sense of energy that spread

Beautifully animated and superbly written, this is a hugely entertaining, frequently

Magical storytelling that makes perfect family entertainment for the Easter holidays

It’s not only better than I thought it was going to be, it’s a lot better than I thought it

Beautifully crafted and effortlessly entertaining, this is an unapologetic triumph.

It respects its audience enough without sticking to cheap shots and bad jokes.

The dragon designs are wonderful, the action is exciting and the anti-worm-rat

How to Train Your Dragon is a visual marvel, and not just because it’s in 3D...

How to Train Your Dragon is a visual marvel, and not just because it’s in 3D...

Though the 3-D effects are awesome, this movie also succeeds in two-dimensions

Indubitably Dreamworks’ best film yet, and quite probably the best dragon movie

0 1 2 3 4 5 6 7 8 9 10

Movie reviewer score indicated by length of bold

How to Train Your Dragon

Everything from the angle of the shot to the speed of the editing projects an end view

Here, Viking life is grim, hostile and heavy with social pressure - kind of like Gossip

The visuals are striking, the script sharp and well-paced and it all wraps up with a bit

Full of wonder, charm and dragons not doubling as stand-up comics.

It’s a brisk, amusing piece that doesn’t have the weary ween that leaves a tinge.

Following a slow, overly verbal start, this dragon tale takes off!

Baruchel, Ferguson and Butler supply a contagious sense of energy that spread

Beautifully animated and superbly written, this is a hugely entertaining, frequently

Magical storytelling that makes perfect family entertainment for the Easter holidays

It’s not only better than I thought it was going to be, it’s a lot better than I thought it

Beautifully crafted and effortlessly entertaining, this is an unapologetic triumph.

It respects its audience enough without sticking to cheap shots and bad jokes.

The dragon designs are wonderful, the action is exciting and the anti-worm-rat

How to Train Your Dragon is a visual marvel, and not just because it’s in 3D...

How to Train Your Dragon is a visual marvel, and not just because it’s in 3D...

Though the 3-D effects are awesome, this movie also succeeds in two-dimensions

Indubitably Dreamworks’ best film yet, and quite probably the best dragon movie

0 1 2 3 4 5 6 7 8 9 10

Movie reviewer score indicated by length of bold

Fig. 8. Movie reviews from Rotten Tomatoes. The length of bold indicates the reviewer’s score.

3.3 Describing Topics

While tag clouds (fig. 1) can possibly show topics in a text, they don’t provide any qualities associated with the topic. In figure 9, characters from Grimm’s Fairy Tales are listed along the left side and adjectives that occur within +/- three words are listed to the right. Adjective frequency is indicated by font weight: one can see that kings are mostly old, princesses are beautiful and girls are little.

Character List of adjectives, weighted by frequency: 2 3 4 5 6 7 9 10

bird beautiful splendid open wooden like hanging

cat little one long

fox old dead young first fast

greel little poor good

hans ill good clear

hansel little like fat

bride false true first right real

king great old one young three angry beautiful married like third ready

princess beautiful young last dear enchanted strange true third free

girl little poor lazy pretty young dead beautiful silly

queen beautiful late little far

wife standing married next poor two beautiful new one dear true first

witch old wicked

Fig. 9. Adjectives associated with characters from Grimm’s Fairy Tales. Heavier weights are more frequent.

3.4 Comparing Countries

A choropleth map is a popular visualization technique that colors each country based on a data value associated with that country. Choropleth maps have problems, such as small countries (e.g. Singapore, Luxembourg) are not visible (fig. 10 top) and only one variable is shown at a time.

Fig. 10. Top: Choropleth map shows HIV prevalence by country color. Bottom: Font weight, caps and italics indicate health expenditures, life expectancy and HIV prevalence respectively.
Instead, the map using three letter country ISO codes can represent multiple variables using font weight, case and italics (figure 10 bottom). Small countries are clearly visible. Complex questions can be answered, such as countries with high spending and short lives (yes, e.g. Rwanda rwa) or low spending and long lives (yes, e.g. Luxembourg, LUX).

3.5 Songs in Text

In written prose, singing may be differentiated from surrounding text, for example, by being set in italics. However, this does not convey any of the song qualities such as the note pitch and note duration. While traditional music notation could be used, this would interrupt the flow of the text and require a lot of space. Instead, syllables could use x-height to indicate note pitch and compressed/expanded fonts to indicate note duration (fig. 11).

![Text Example](image)

Fig. 11. Song text with note pitch indicated by x-height (i.e. high notes have a tall x-height, e.g. vous) and note duration indicated by font width (i.e. short notes are narrow, e.g. son-nez; and long notes are expanded, e.g. dong).

4 Conclusion

The examples here suggest that there is a wide range of possibilities for unique visualizations using font attributes. Examples shown cross many domains. Many of the examples use font weight and other font attributes are under-explored. A few additional examples can be found in [22,23,24].

Some interesting implications arise from this work that may be of interest for further investigation by type designers. One example uses variable x-height: traditionally the x-height is an inherent property of a typeface. Is it feasible to have a type family with multiple x-heights?

Many of the examples use multiple font weights. These are largely restricted to sans serif fonts which currently provide the greatest variation in weights. Can a serif typeface be provided in a broad variety of weights? Is it feasible to design a lightweight blackletter or a black cursive?

Many of the examples use multiple font attributes within a single visualization, for example combining variations across weight, italics, case; or combining variations in x-height and font widths. Can type families be created that provide even more variation across attributes, for example, a typeface with multiple weights and multiple oblique angles? Given potentially a resulting very large type family - how is it installed and licensed?

Attempting to mix and match different typefaces to encode different categories can be difficult. For example using a serif, sans serif and a slab serif is difficult when it is desired to have matching weights, oblique angles and so forth.

From a data visualization perspective, while there are many examples shown here - these are all speculative applications and have not been tested with real users. How effective these techniques are has yet to be established. However, the breadth of examples suggest rich possibilities for future work.

References