

Considerations in designing linguistic visualizations

Chris Culy
chrisculy.net



AVML Workshop
26 September 2016

Three distinct goals of visualizations

- Exploration
- Comprehension (analysis)
- Communication (presentation)

A visualization may be suitable for some goals but not others

Reference: B.H. McCormick, T.A. DeFanti, M.D. Brown. 1987. Visualization in Scientific Computing, Computer Graphics 21:6. ACM SIGGRAPH.

Three evaluations of visualizations

- As a type: are they fit for purpose?
- Of instances: are they fair or misleading?
 - Lies, damned lies, and ... visualizations?
- For you: do they do what you need?
cf. Goals: exploration, comprehension, communication

What makes L/L data special?

1. Language is not *mappable*
2. Individual pieces of L/L data are *meaningful*
3. Basic L/L data is often *computed*, not observed

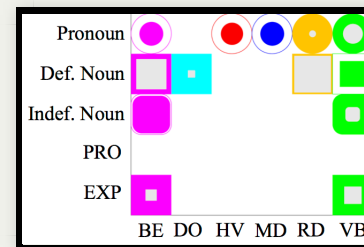
Charts

- Charts are basic, but important
 - interactive updating is useful for exploration
- Lots of tools to make charts. An interesting one is Tableau. And some others: Spotfire, Qlik; Raw, Vega, Lyra;
- A good comparison of chart types with the same data is on the Protovis site

References: Tableau, Spotfire, Qlik, Raw, Vega (and Lyra)

Elaborating charts

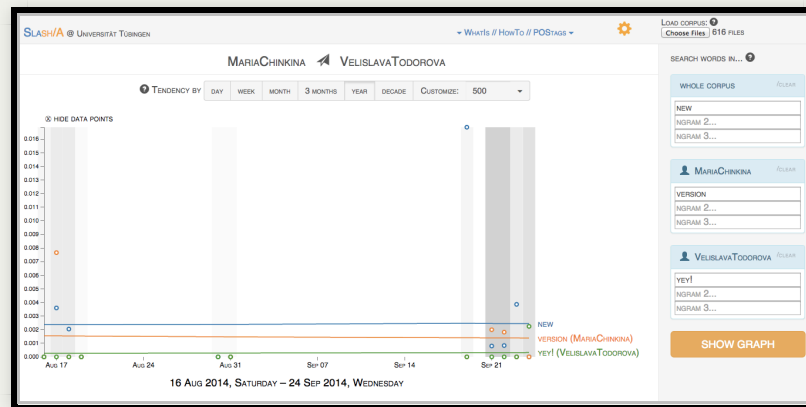
Icelandic Diachronic Corpus



Reference:

Butt, Miriam, Bögel, Tina, Kotcheva, Kristina, Schätzle, Christin, Rohrdantz, Christian, Sacha, Dominik, Dehé, Nicole and Daniel Keim. 2014. 'V1 in Icelandic: A Multifactorial Visualization of Historical Data'. Proceedings of the LREC 2014 Workshop on Visualization as added value in the development, use and evaluation of LRs (VisLR). Reykjavik, Iceland.

Connecting charts with the source data



[demo]

Source: [Slash/A](#) by Slava Todorova and Maria Chinkina

Design principles and tips

Gestalt Principles

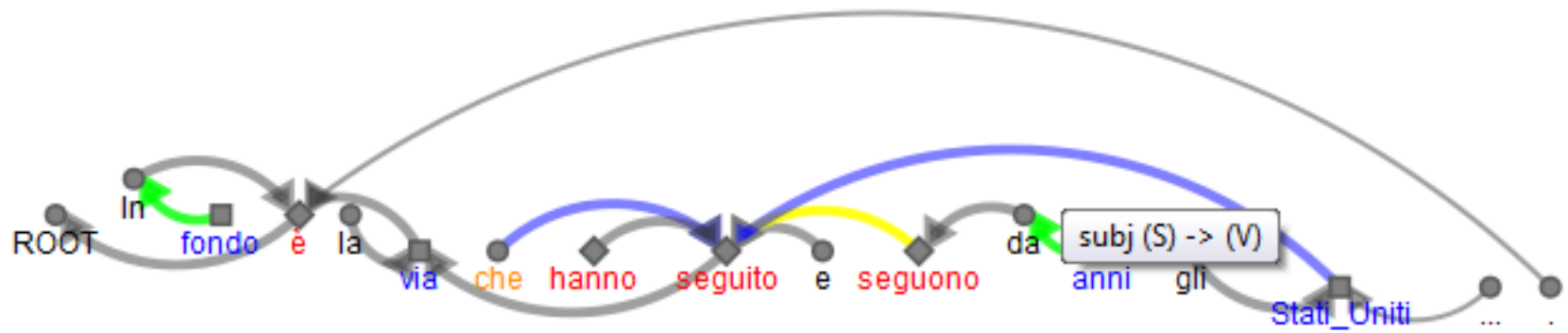
We *tend* to perceive things that are similar as belonging to a common group. Similar by:

- Proximity
- Shape, size, texture, or color
- Connected by lines
- And more ...

Watch out for inappropriate Gestalt effects.

References:

C. Ware. 2012. Information Visualization: Perception for Design. Chapter 6.



“Pre-attentive” visual features

Certain visual aspects are very prominent

- “Low level” visual features are processed sooner and faster than higher level patterns
 - Color, grey, shape, size, length, orientation, motion, (and others)
- Recognition is facilitated (biased) by higher level task and attention

References:

C. Ware. 2010. Visual Thinking for Design. Chapter 2

Two uses of visual features

- Highlighting
- Encoding information = visual variables (cf. Gestalt)

References:

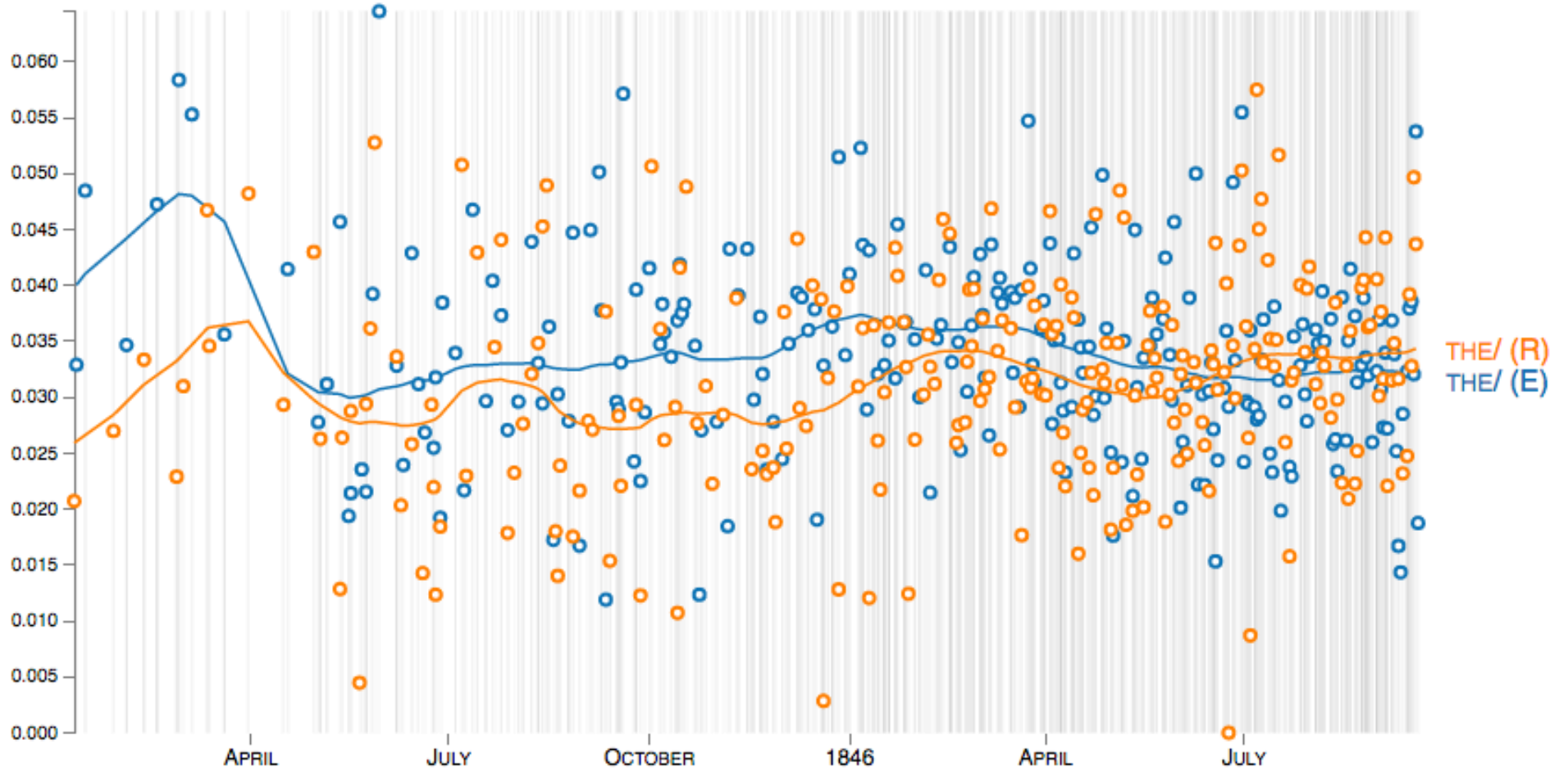
C. Ware. 2010. Visual Thinking for Design. Chapter 2

Design: Some visual variables and their limits

- Position – most flexible
- Color – 5-10
- Size – grouping ~5; distinguishing ~20
- Brightness – grouping < 7; distinguishing ~ 10
- Shape – easier when fewer, not close together
- Pattern – easier when fewer, not close together

It is difficult to encode many values with a visual variable.

Reference: S. Carpendale. 2003. Considering Visual Variables as a Basis for Visualization. Research report 2001-693-16, Department of Computer Science, University of Calgary.



Color is powerful, but be careful

- Be aware of constraints of colorblindness or low-vision
 - Allowing users to set colors / color schemes is the most flexible. **Scheme 2** **Scheme 3**
 - Have enough contrast between text and background
- Hue is unordered, but can be given (cultural) orderings (e.g. temperature, elevation)
- For ordered data, darker / more saturated for higher values
- Lighter, less saturated colors for larger areas
- Darker, more saturated colors for smaller areas and lines

Resources: [Color Brewer](#), [Sim Daltonism](#): OS X color blindness simulator (there are others), [WebAIM](#) contrast and accessibility checker

How should visualizations interact with the user?

Shneiderman's "Information Seeking Mantra"

1. Overview first
2. Zoom and filter
3. Details on demand

References: B. Shneiderman. 1996. The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of the 1996 IEEE Symposium on Visual Languages, VL '96.

A few interaction techniques

- Semantic Zoom
- Tooltips
- Coordinated views

Some additional considerations

- Ability to see the original data
- Ability to navigate backwards and forwards between states (not common)
- Ability to link to a state (not common)
- Ability to annotate the visualization (very rare)

References: IBM's [original version of Many Eyes](#)

J.S. Yi, Y.A. Kange, J.T. Sasko, J.A. Jacko. 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*. 13:6

J. Unsworth. 2000. "Scholarly Primitives: what methods do humanities researchers have in common, and how might our tools reflect this?" Symposium on Humanities Computing: formal methods, experimental practice. King's College, London, May 13, 2000.

Further issues

- What are the tasks people are doing? (More tomorrow)
- How important are data uncertainty and data errors to the users?
- How can we visualize confidence/error information, if we have it?
- How can we incorporate peoples' preferences?

More of my resources

- These slides (and tomorrows) are on my web site chrisculy.net/lx/presentations/
- My workshop for the British Association of Applied Linguistics (BAAL)
- My ESLLI course (with Miriam Butt): Visualization theory

Blue skies and green pastures

What would you do with an unlimited budget?

Thank you

Chris Culy
chrisculy.net

