

# Imagery Lenses for Visualizing Text Corpora

**Min Chen**

Professor of Scientific Visualization

Think New Shapes  
*Rado (2004)*



THINK NEW SHAPES.



Oxford e-Research Centre  
University of Oxford  
[min.chen@oerc.ox.ac.uk](mailto:min.chen@oerc.ox.ac.uk)

JISC DID Programme Meeting,  
Bristol, 8 November, 2012

## Utah Members of the Project

- Dr. Julie Lein
  - *RA, English literature*
- Professor Katharine Coles
  - *PI, poet, English literature*
- Professor Chris Johnson
  - *CI, numerical analysis, scientific visualization*
- Dr. Miriah Meyer
  - *“volunteer CI”, bio-information visualization*

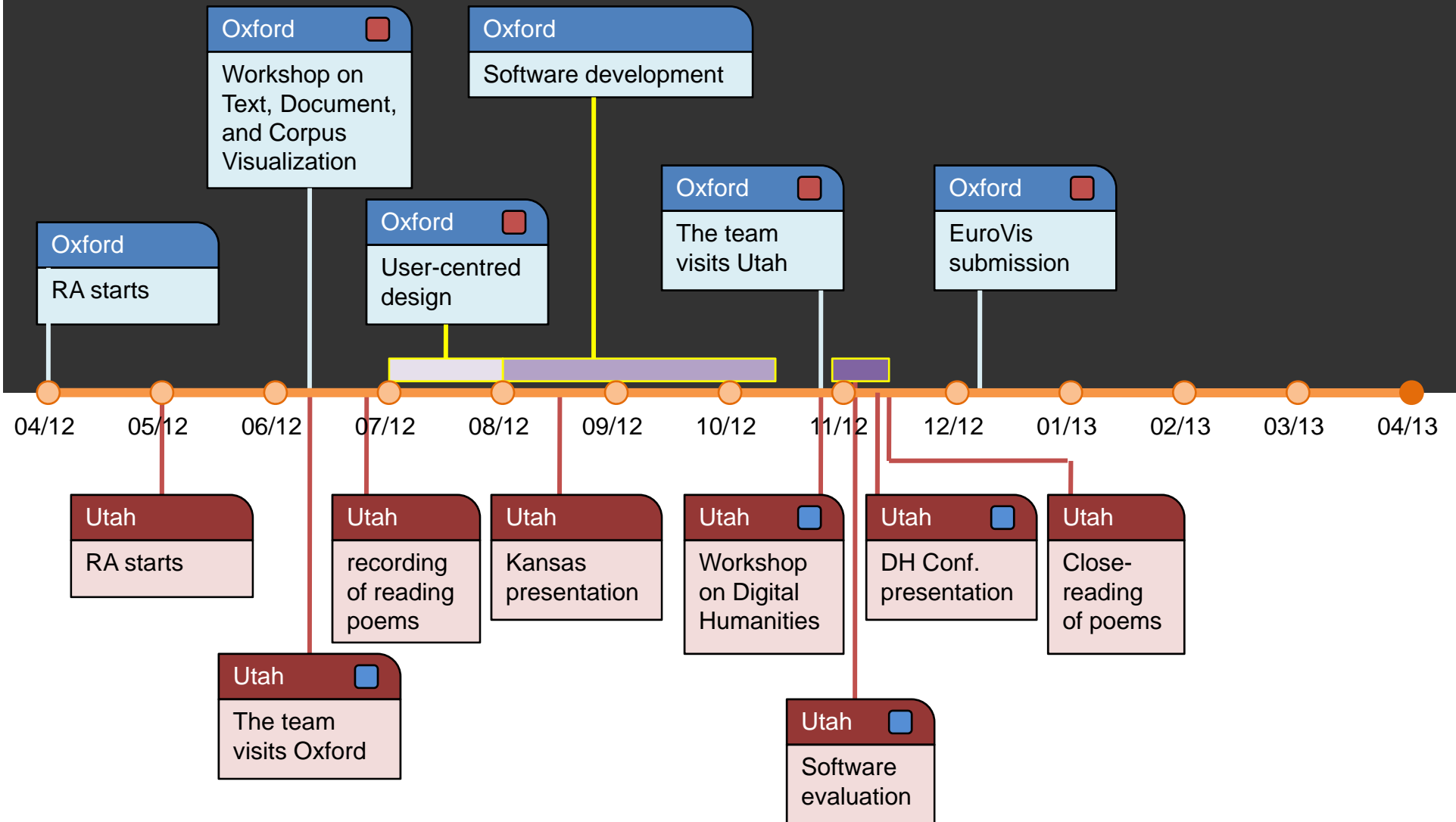


## Oxford Members of the Project

- Dr. Alfie Abdul-Rahman
  - *RA, visualization*
- Professor Min Chen
  - *PI, visualization*
- Dr. Martin Wynne
  - *CI, linguistics*
- Professor Anne Trefethen
  - *CI, numerical analysis*
- Eamonn Maquire
  - *“volunteer RA”, bio-information visualization*



# Summary of Our Progress To Date



3d analysis applications biomedical  
 comparative context continuous curve data

design diagrams dimension direct efficient  
 exploration fast field flow framework  
 graphic high illustrative image

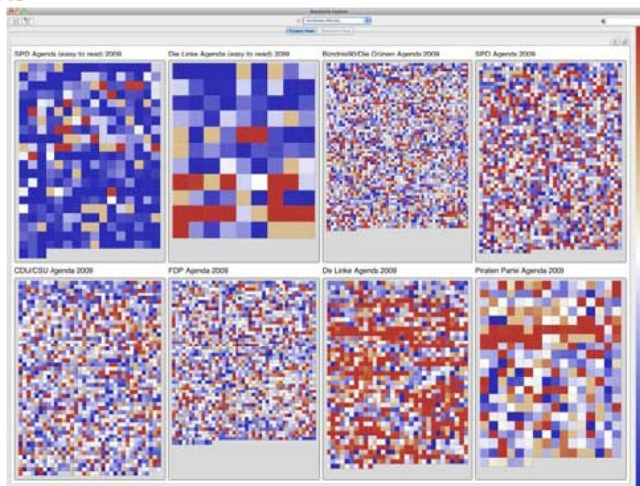
information interactive

local mapping model multidimens  
 perception projection quality ray  
 rendering segmentation simula  
 space surfaces system theory tir  
 uncertainty using vector view

visualizati

volume

by Oelke, Spretke, Stoffel,  
 Keim, IEEE VAST 2010

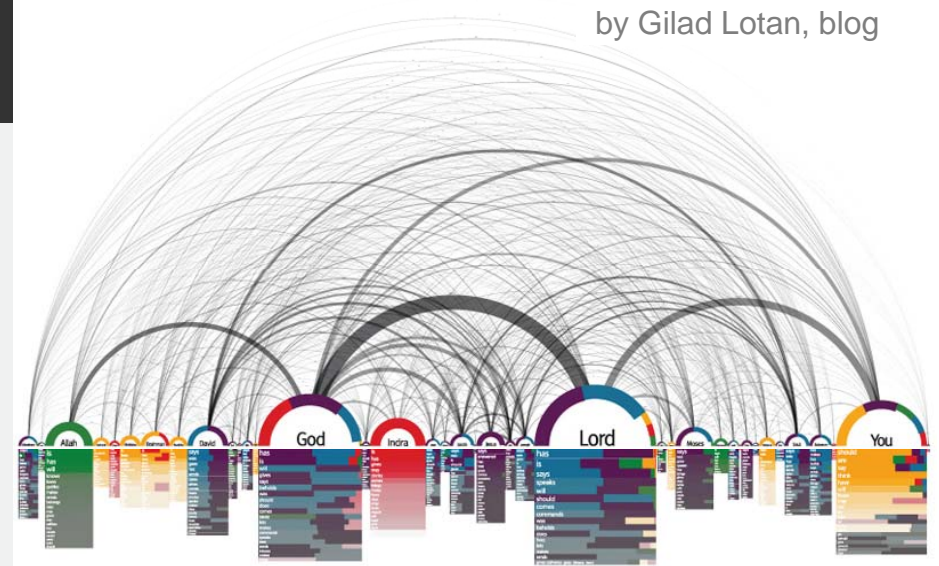


(b) Feature: Vocabulary Difficulty

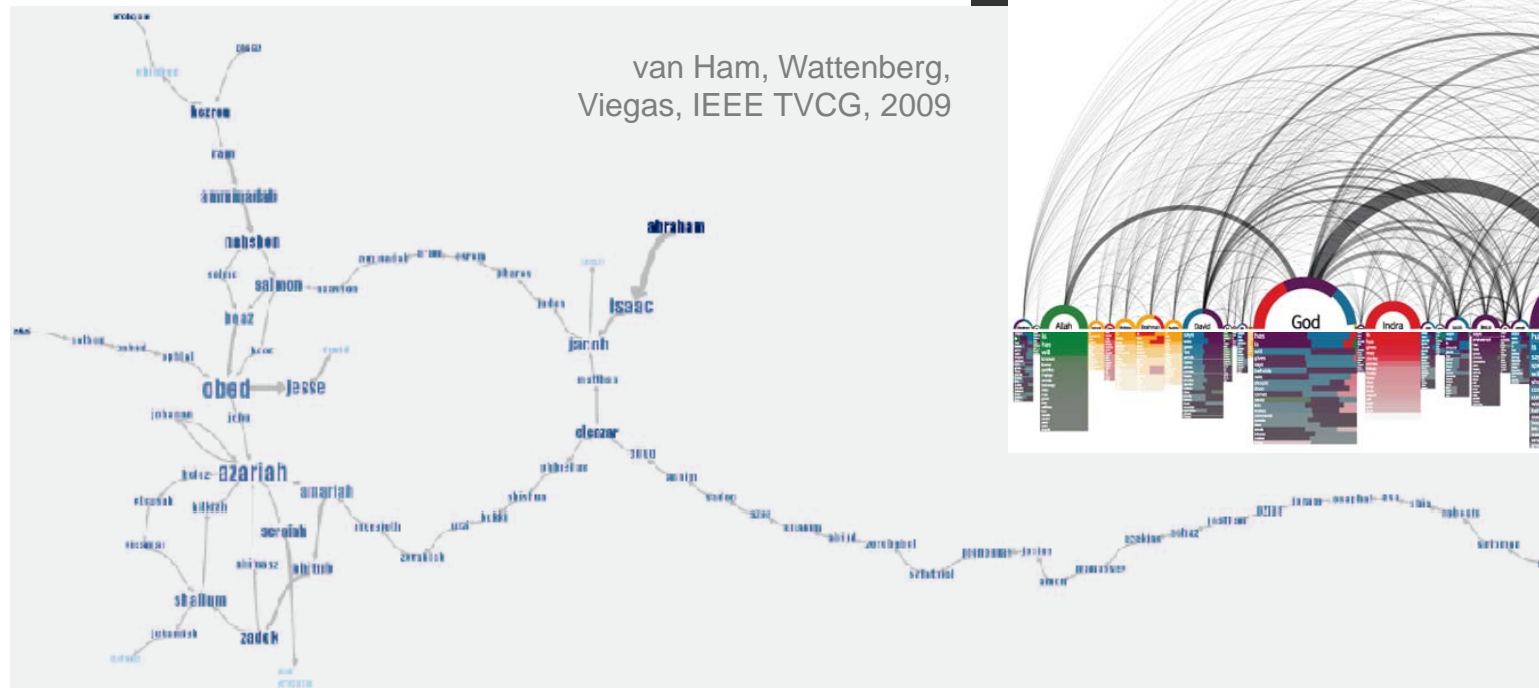
# State of the art

- Tag clouds
- Pixel-based visualization
- Word networks

by Gilad Lotan, blog



van Ham, Wattenberg,  
 Viegas, IEEE TVCG, 2009



# Night by Louise Bogan (Sound Components)

View:  IPA  Text

Legend: ■ Vowel ■ Consonant ■ Punctuation

ðə kɔld rɛmɔt ədʒlɛndz  
The cold remote islands

ænd ðə blu ɛstʃuəri:z  
And the blue estuaries

wɛr wət brɪðz, brɪðz  
Where what breathes, breathes

ðə rɛstlɛs wɪnd əv ðə ɪnlɛts,  
The restless wind of the inlets,

ænd wət drɪŋks, drɪŋks  
And what drinks, drinks

ðə ɪnkɛmɪŋ tɑ:d;  
The incoming tide;

wɛr ʃɛl ænd wɪd  
Where shell and weed

wɛt əpɑn ðə sɔlt wɑʃ əv ðə si:  
Wait upon the salt wash of the sea,

ænd ðə klɪr naɪts əv stɑrz  
And the clear nights of stars

swɪŋ ðɛr laɪts wɛstwɜrd  
Swing their lights westward

tʊ sɛt bɛhɑnd ðə lænd;  
To set behind the land;

wɛr pɛls klɪŋɪŋ tʊ ðə rɔks  
Where pulse clinging to the rocks

rɛnu:z ɛtsɛɪf fɛrɛvər;  
Renews itself forever;

wɛr, əgɛn ɑn klɔdɪs naɪts,  
Where, again on cloudless nights,

ðə wɔtər rɛflɛkts  
The water reflects

ðə fɜrmɛmɛnt's pɑrʃəl sɛtɪŋ;  
The firmament's partial setting;

-o rɛmɛmbər  
-O remember

ɪn jɔr nɛrɔɪŋ dɑrk ɑwɜrz  
In your narrowing dark hours

ðæt mɔr θɪŋz muv  
That more things move

ðæn blʊd ɪn ðə hɑrt.  
Than blood in the heart

## A Basic Visualization

- Spatial structure
- Text
- Phonetic symbols
  - *interactional phonetic association*
- Basic classification
  - *vowel, consonant, and punctuation*

### Basic word

Latin alphabet:

Latin type:

### Basic phonetic

Phonetic alphabet:

Phonetic type:

Vowel position:

Consonant type:

Consonant features:

Stress:

Vowel length:

Syllable break:

### Word action

Word class:

Word profile:

Repeated word:

### Phonetic action

Rhyme:

Assonance:

Consonance:

Alliteration:

End rhyme:

Internal rhyme:

Rhymed stanza break:

Rhymed line break:

Same rhyme spelling:

Semantic connection:

### Poem action

Structure break:

Sound frequency:

Rhyme frequency:

Word frequency:

- Colour of arc
- Texture of arc
- Arc styles
- Thickness of arc
- Shading of arc
- Transparency of arc
- Height of arc
- Different types of shape
- Different types of symbol
- Different types of character
- Colour of shape
- Colour of symbol
- Colour of character
- Size of shape
- Size of symbol
- Size of character
- Orientation of shape
- Orientation of symbol
- Orientation of character

# Parameters of Poems

- Some 30 parameters

- What visual channels can be used?

- Colour



- Shape



- Size



- Orientation

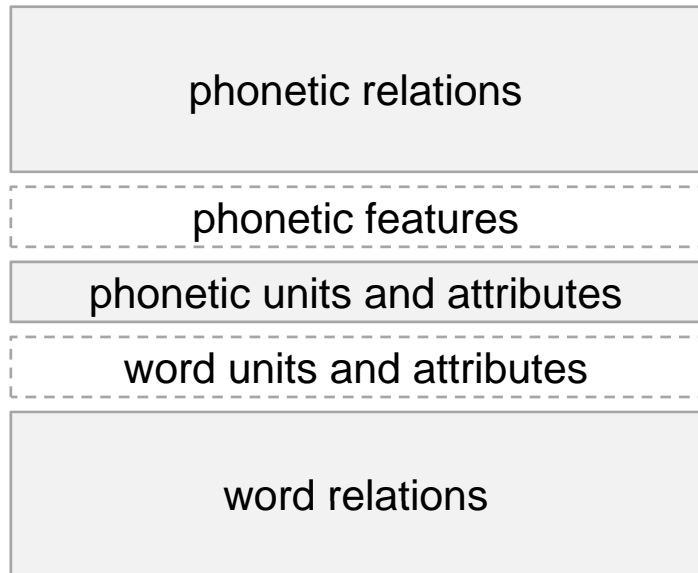
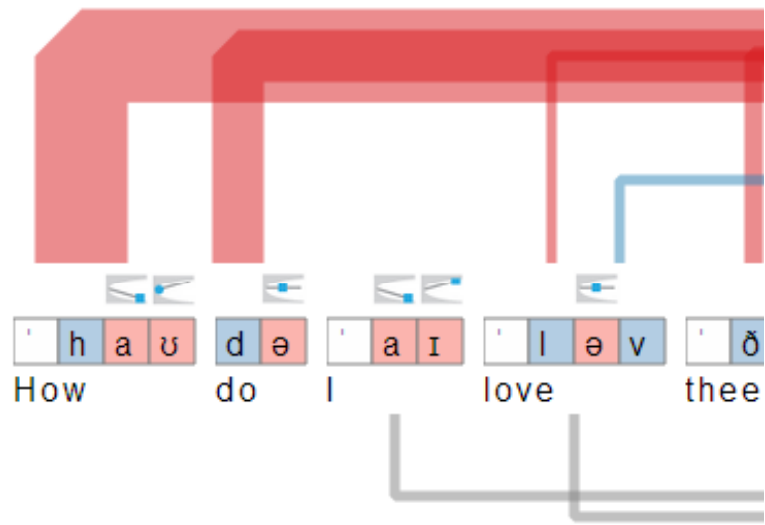


- Connection



- ....





## Rule-based selection

- Spatial separation
  - *semantic categories correlate to regions in visualization*
- Channel type matching
  - *nominal*
  - *ordinal*
  - *interval*
  - *ratio*
  - *relational*
- Channel capacity
  - *minimal mapping range*
  - *ideal mapping range*
- Already-in-use

# Some of the Current Visualization

## Sonnets from the Portuguese 43 by Elizabeth Barrett Browning



[About](#) [Contact](#)

Poem: Sonnets from the Portuguese 43 - Elizabeth Barrett Browning

### Rule Mappings

#### Basic phonetic

Variable	Region	Mapping	View	Legend
Vowel position:	Additional phonetic	Pictogram	<input checked="" type="checkbox"/>	Rounded:  Close-front  Close-central  Close-back  Mid-front  Mid-central  Mid-back  Open-front  Open-central  Open-back Unrounded:  Close-front  Close-central  Close-back  Mid-front  Mid-central  Mid-back  Open-front  Open-central  Open-back
Phonetic type:	Phonetic	Colour of background	<input checked="" type="checkbox"/>	<span style="color: red;">■</span> Vowel <span style="color: blue;">■</span> Consonant <span style="color: gray;">■</span> Punctuation
Phonetic alphabet:	Phonetic	Different types of symbol	<input checked="" type="checkbox"/>	
Stress:	Phonetic	Colour of symbol	<input checked="" type="checkbox"/>	<input type="checkbox"/> Primary stress <input type="checkbox"/> Secondary stress
Vowel length:	Phonetic	Size of symbol	<input checked="" type="checkbox"/>	<input type="checkbox"/> Long <input type="checkbox"/> Half-long <input type="checkbox"/> Extra-short
Syllable break:	Phonetic	Transparency of background	<input checked="" type="checkbox"/>	<input type="checkbox"/> Syllable break

#### Basic word

Variable	Region	Mapping	View
Latin alphabet:	Word	Different types of character	<input checked="" type="checkbox"/>

#### Poem action

Variable	Region	Mapping	View	Legend
Structure break:	Phonetic	Size of shape	<input checked="" type="checkbox"/>	Line break

#### Phonetic action

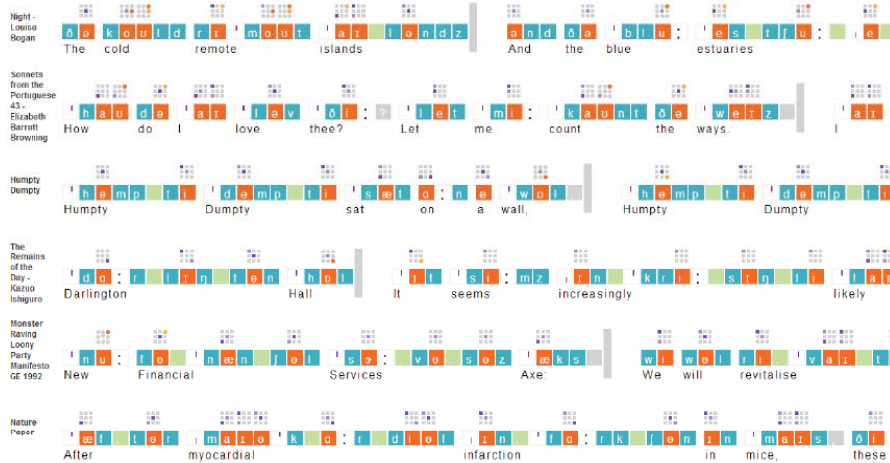
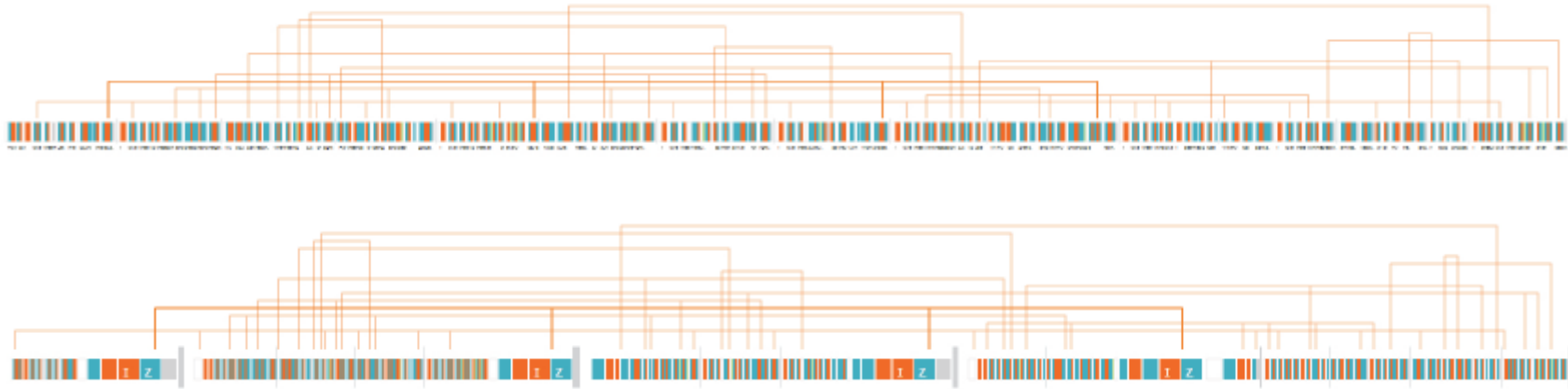
Variable	Region	Mapping	View	Legend
Type of relation:	Phonetic connection	Colour of arc	<input checked="" type="checkbox"/>	<span style="color: blue;">—</span> Rhyme <span style="color: red;">—</span> Alliteration
Rhyme:	Phonetic connection	Height of arc	<input checked="" type="checkbox"/>	
Alliteration:	Phonetic connection	Thickness of arc	<input checked="" type="checkbox"/>	

#### Word action

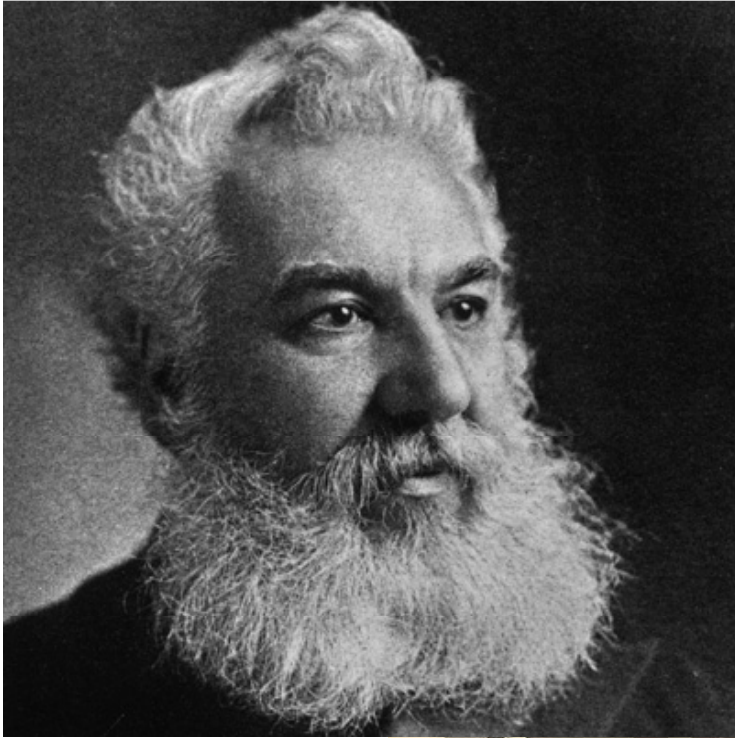
Variable	Region	Mapping	View	Legend
Repeated word:	Word connection	Height of arc	<input checked="" type="checkbox"/>	<span style="color: gray;">—</span> Repeated word



# Some of the Current Visualizations



Alexander Graham Bell (1847-1922)



## Telephone

- In the 1870s, Bell travelled around to give demos 'in concert halls, where full orchestras and choruses played "America" and "Auld Lnag Syne" into his gadgetry.'
- Around 1880, Queen Victoria installed a pair of telephones at Winsor and Buckingham Palace

“Mr. Watson, come here. I want to see you”, said Bell

“Mr. Information, come here.  
I want to see you.”