EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

PHILOSOPHISCHE
FAKULTÄT
Seminar für Sprachwissenschaft

Visualization of Language and Linguistic Information

Chris Culy

Arbeitstreffen: Netzwerk Lexicographie

IDS Mannheim

20 November 2013

Original at: http://www.sfs.uni-tuebingen.de/~cculy/presentations/NetzwerkLexicographie2013/index.html

My interest in language and visualization

Mali

Bambara: malonyinina o malonyinina
Fulfulde: mi yi'ii ɗum e giɗum
Donno Sɔ Dogon: Oumar Anta inyemɛñ waa be gi

C. Culy. 1988. Basic Donno So (Dogon), Grammar notes and glossary written for Peace Corps-Mali.

Source: OpenStreetMap

My interest in language and visualization

Takelma texts

Mi^{i} aga sgísi ā′k!à da^{ɛ}ána-u wa-iwi^{i} dálhiwilīk‘^{w}. Ganēhi^{ɛ} bo^{u} nēxada^{ɛ} “Wa-iwí^{i} di eīt‘? Wa-iwí^{i} mī′^{ɛ}wa,” nagá-ihìs; sgísi^{ɛ}a mi^{i} gelwaínia gelgulùk‘^{w}. Ganēhi^{ɛ} ánī^{ɛ} t‘ayàk‘ gwī^{ɛ}néi hawúxda^{a}. “K‘ádi gi^{iɛ}à? K‘a-ilā′p‘a mī^{ɛ}wa nagásbi^{ɛ}n,” nagáhi^{ɛ}. Sgísi lap‘ãm xamgwidìk‘^{w}. “Ma dí k‘ai^{ɛ}lā′p‘a yuda^{ɛ}? lap‘a^{a}m nánsbina^{ɛ},” nagáhi^{ɛ} lap‘ãm. Gé de^{ɛ}winít‘hì. Gweldi; ba^{a}bi^{ɛ}t‘ lé^{e}p‘lap‘.

Now those just scattered off, Grizzly Bear did chase the people around. Now this Coyote, for his part, did run off with the chieftainess girl. Then, 'tis said, after a little while, "Are you a female? It must be a female," he thought; Coyote now, for his part, did wish to sleep with her. Tunc nihil vulvae repperit. "What did I, for my part, (take)? That you were a woman I thought," he said to her. Coyote threw Frog into the water. "Do you think you will be a woman? Frog you will always be called," he said to Frog. Proceeding just up to there (it goes). 'Tis finished. Go gather and eat your ba|ap'-seeds.

Source: Takelma Texts

My interest in language and visualization

Recipes, letters

Bake Ø until done

EBB: Overjoyed I am

My interest in language and visualization

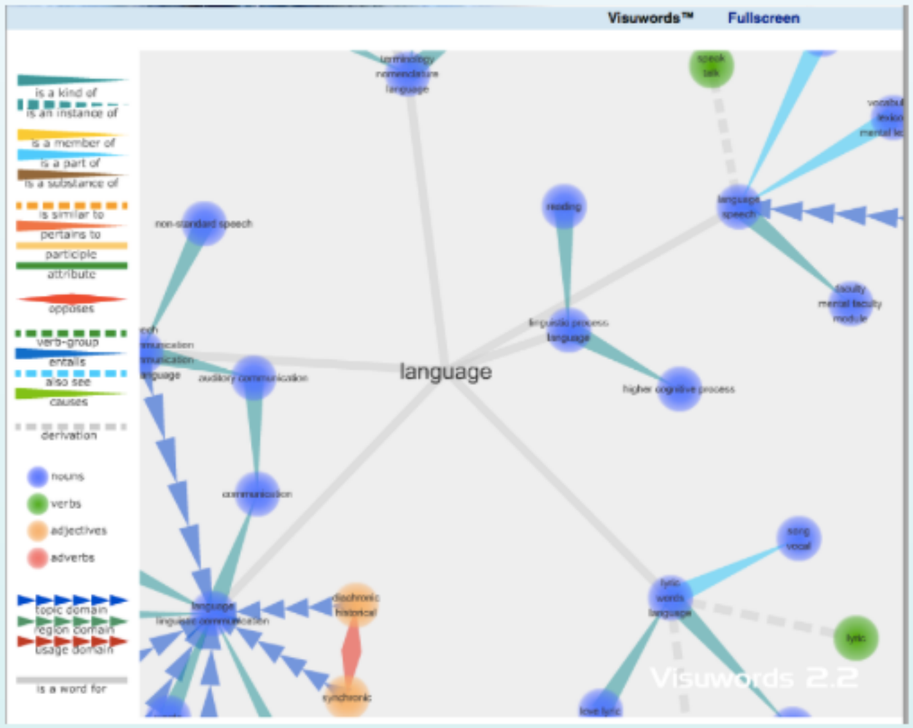
Visualizations as tools

Source: C. Culy, E. Chiocchetti, and N. Ralli. 2013

Visualizations put ideas into our heads ...

3 Separate goals of visualizations

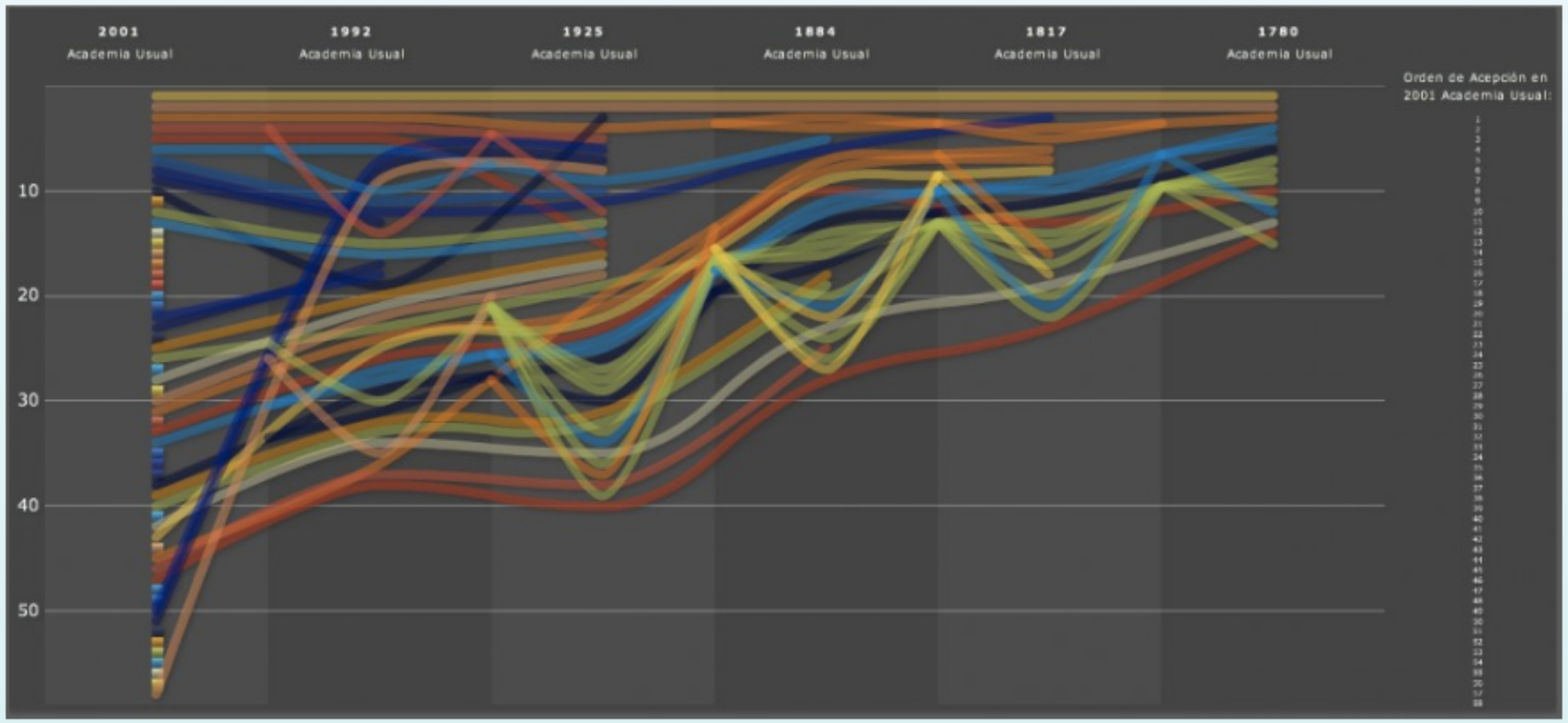
1. Exploration



Source: Visuwords™

3 Separate goals of visualizations

2. Comprehension (Analysis)



Source: R. Theron and L. Fontanillo. 2013

3 Separate goals of visualizations

3. Communication (Presentation)

Source: C. Collins, S. Carpendale, and G. Penn. 2009

Research (and practical) question

- **Which visualizations are best suited for which of these goals?**

Finding the right visualization

1. Grinstein's Grand Challenge
2. Pre-existing first steps

   Tableau, Spotfire

3. **Limitations**
   - Limited types of data: numbers, dates, geographic, categories
   - No notion of task
   - No notion of preferences

Sources: Georges Grinstein, Tableau, Spotfire

Tackling the Grand Challenge

Dataset genres

Person-oriented correspondence

(What *other* (lexicographic) dataset genres do we have?)

Sources: on archive.org, on archive.org

Language is different

1. Language is not *mappable*
2. Individual pieces of data are meaningful
3. **Much *linguistic* data is computed, not observed**

Rethinking language visualizations

Num	Date	From	-5	-4	-3	-2	-1	hit	+1	+2	+3	+4	+5
1	1845-01-10	R	Hatcham/NP	,/,	Surrey/NP	./SENT	I/PP	**love/VBP**	your/PP$	verses/NNS	with/IN	all/PDT	my/PP$
1	1845-01-10	R	,/,	as/IN	I/PP	say/VBP	,/,	**love/VBP**	these/DT	books/NNS	with/IN	all/PDT	my/PP$
1	1845-01-10	R	my/PP$	heart/NN	--/:	and/CC	I/PP	**love/VBP**	you/PP	too/RB	./SENT	Do/VBP	you/PP
5	1845-01-28	R	and/CC	dissertate/VB	upon/RP	that/IN	I/PP	**love/VBP**	most/JJS	and/CC	least/JJS	--/:	I/PP
7	1845-02-11	R	so/RB	,/,	if/IN	"/``	I/PP	**love/VBP**	you/PP	"/``	were/VBD	always/RB	outspoken/JJ
8	1845-02-17	E	not/RB	see/VB	where/WRB	./SENT	I/PP	**love/VBP**	the/DT	drama/NN	too/RB	./SENT	I/PP
8	1845-02-17	E	princes/NNS	in/IN	poetry/NN	./SENT	I/PP	**love/VBP**	them/PP	through/IN	all/PDT	the/DT	deeps/NNS
9	1845-02-26	R	"/"	so/RB	thoroughly/RB	does/VBZ	he/PP	**love/VBP**	and/CC	live/VBP	by/IN	it/PP	./SENT
13	1845-03-12	R	properly/RB	for/IN	it/PP	,/,	I/PP	**love/VBP**	and/CC	wish/VBP	you/PP	well/RB	!/SENT
17	1845-04-18	E	you/PP	--/:	and/CC	whatever/WDT	you/PP	**love/VBP**	or/CC	hate/VBP	,/,	whatever/WDT	charms/VBZ
19	1845-05-02	E	Mr./NP	Browning/NP	,/,	that/WDT	we/PP	**love/VBP**	the/DT	darkness/NN	and/CC	use/VB	a/DT
30	1845-05-24	E	of/IN	my/PP$	aunts/NNS	whom/WP	I/PP	**love/VBP**	,/,	and/CC	have/VBP	not/RB	met/VBN
32	1845-05-25	E	every/DT	word/NN	./SENT	"/``	I/PP	**love/VBP**	the/DT	truth/NN	and/CC	can/MD	bear/VB
59	1845-07-09	R	And/CC	I/PP	,/,	too/RB	,/,	**love/VBP**	to/TO	have/VB	few/JJ	friends/NNS	,/,
87	1845-08-25	E	I/PP	,/,	for/IN	one/CD	,/,	**love/VBP**	him/PP	!/SENT	and/CC	when/WRB	,/,
102	1845-09-13	R	as/IN	I/PP	observed/VBD	,/,	I/PP	**love/VBP**	you/PP	as/IN	you/PP	now/RB	are/VBP
126	1845-10-13	R	bless/VBP	you/PP	and/CC	all/RB	you/PP	**love/VBP**	!/SENT	dearest/JJS	,/,	I/PP	am/VBP
128	1845-10-15	R	,/,	my/PP$	own/JJ	,/,	dearest/JJS	**love/VBP**	,/,	that/IN	this/DT	is/VBZ	for/IN
136	1845-10-23	R	in/IN	the/DT	least/JJS	./SENT	I/PP	**love/VBP**	you/PP	because/IN	I/PP	love/VBP	you/PP
136	1845-10-23	R	I/PP	love/VBP	you/PP	because/IN	I/PP	**love/VBP**	you/PP	;/:	I/PP	see/VBP	you/PP
136	1845-10-23	R	,/,	live/VBP	my/PP$	life/NN	,/,	**love/VBP**	my/PP$	love/NN	./SENT	When/WRB	I/PP
138	1845-10-25	E	loved/VBD	him/PP	tenderly/RB	(/(	and/CC	**love/VBP**	him/PP	)/)	,/,	.../:	and/CC
144	1845-11-04	R	love/NN	;/:	not/RB	as/IN	I/PP	**love/VBP**	you/PP	--/:	not/RB	for/IN	--/:
166	1845-12-04	E	if/IN	I/PP	feel/VBP	that/IN	you/PP	**love/VBP**	me/PP	,/,	can/MD	I/PP	help/VB
169	1845-12-08	E	ripen/VB	the/DT	knowledge/NN	./SENT	They/PP	**love/VBP**	Tennyson/NP	so/RB	much/RB	that/IN	the/DT
177	1845-12-19	R	to/TO	avoid/VB	writing/VBG	that/IN	I/PP	**love/VBP**	and/CC	love/VBP	and/CC	love/VBP	again/RB

Rethinking language visualizations

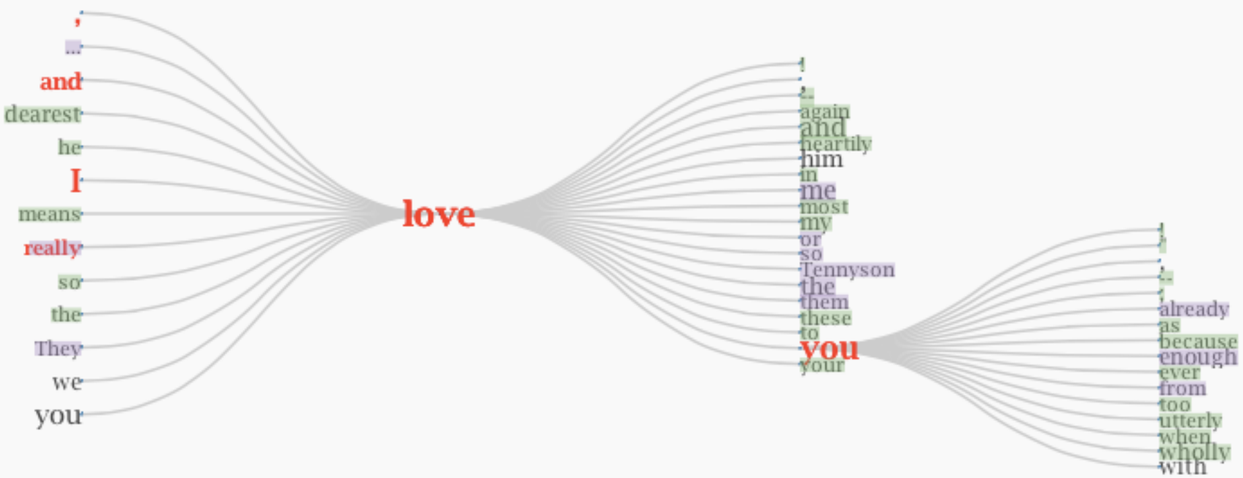
Input file: Choose File no file selected Word to use: love/VBP

Sort branches by: token POS branching

Limit first left POS to:

Limit first right POS to: Update filters

Find in tree: [aA] Clear Show POS



Uses: DoubleTreeJS

Concerning data

(What types of data are especially relevant?)

- What specialized visualizations are especially relevant to the data?

(How important are data uncertainty and data errors?)

- What should we do in the visualization about uncertainty in the data?

- What kinds of mismatches are there between the original data models and the visualization data models?

An aside: Challenge!

11. *To sing*, שִׁיר *shir*, 3*a*.
Zeph. 2. 14 (their) voice shall sing in the windows; de.
12. *A song*, שִׁירָה *shirah*.
Isa. 23. 15 after the end . . shall Tyre sing as an harlot
13. *To sing an ode*, ᾄδω *adō*.
Eph. 5. 19 singing and making melody in your heart
Col. 3. 16 singing with grace in your hearts to the L.
Rev. 5. 9 they sung a new song, saying, Thou art
14. 3 they sung as it were a new song before the
15. 3 they sing the song of Moses the servant of
14. *To sing praise with a musical instrument*, ψάλλω.
Rom. 15. 9 I will confess . . and sing unto thy name
1 Co. 14. 15 will sing with the spirit, and I will sing
SING for joy, to cause to —
To cause to sing or cry aloud, רָנַן *ranan*, 5.
Job 29. 13 I caused the widow's heart to sing for j.

after the end ~~of seventy years~~ shall Tyre sing as an harlot.
For this cause I will confess ~~to thee among the Gentiles,~~ and sing unto thy name.
I caused the widow's heart to sing for j~~oy~~.

Source: Young's Concordance to the Bible

Typical tasks

Task	Proposed by			
	Shneiderman	**Keim**	**Yi**	***Unsworth***
Overview	✓			
Zoom, Abstract/Elaborate	✓	✓	✓	
Details-on-demand	✓	✓		
Filter, Select, *Selection*	✓	✓	✓	✓
Relate, Connect, *Comparing*	✓	✓	✓	✓
Extract, *Sampling*	✓			✓
Explore, *Discovering*			✓	✓
Reconfigure			✓	
Encode, *Representing*			✓	✓
History	✓			
Annotating				✓
Referring, linking				✓

Sources: Shneiderman 1996, Keim et al. 2006, Yi et al. 2007, Unsworth 2000; Many Eyes

A question about tasks

(What other relevant tasks are there?)

e.g.

- Find typical/unusual data
- Find connections within the data
- Find connections to external information
- Data modification

References: Bamman et al. 2007, Passarotti 2013

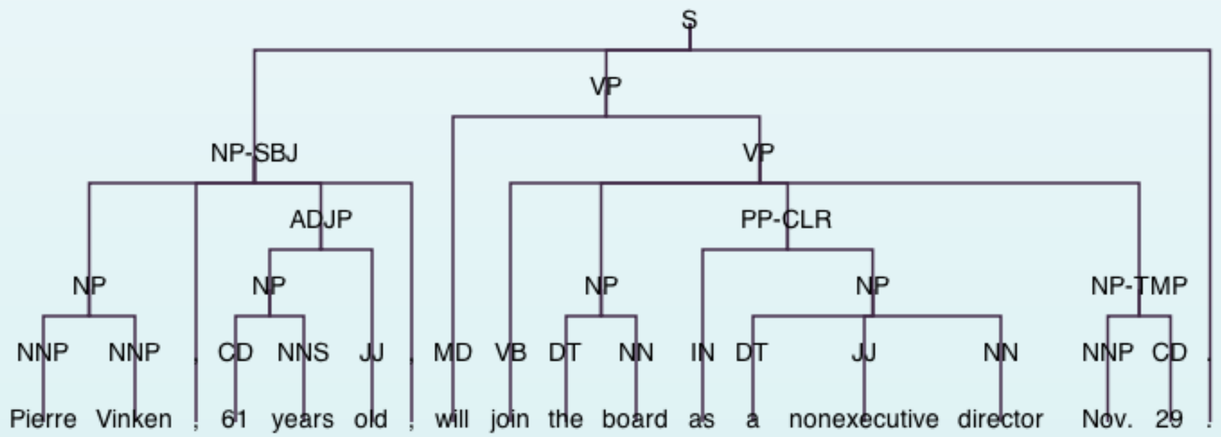
Visualizations for tasks and users

[S[NP-SBJ[NP[NNP[Pierre]][NNP[Vinken]]][,[,]][ADJP[NP[CD[61]][NNS[years]]][JJ[old]]][,[,]][VP[MD[will]][VP[VB[join]][NP[DT[the]][NN[board]]][PP-CLR[IN[as]][NP[DT[a]][JJ[nonexecutive]][NN[director]]]][NP-TMP[NNP[Nov.]][CD[29]]]][.

Draw!

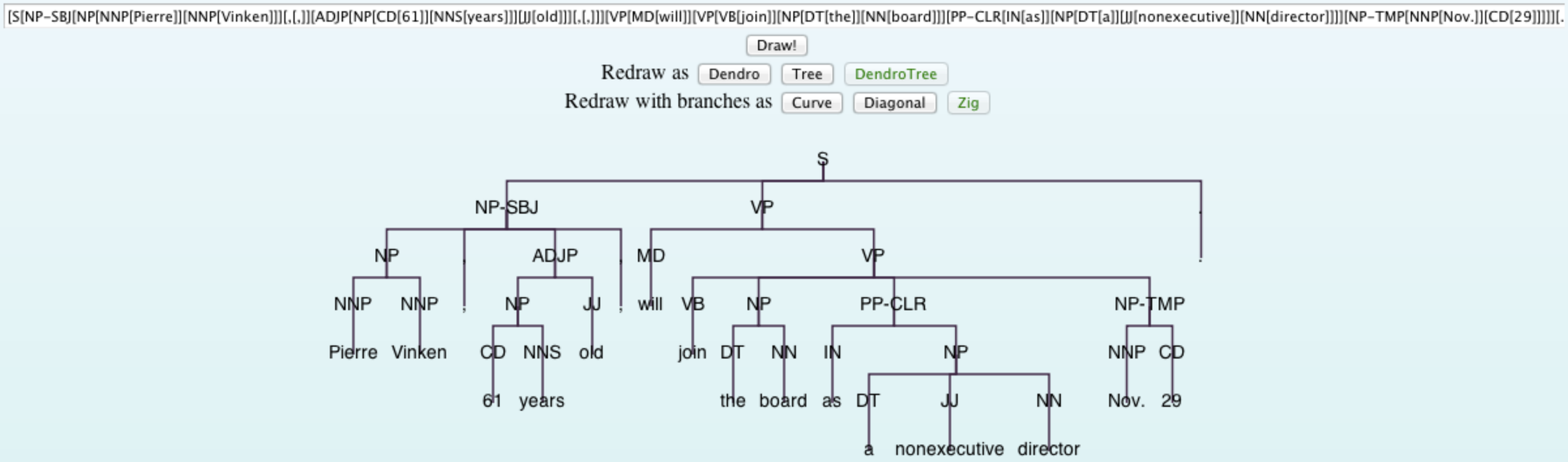
Redraw as Dendro Tree DendroTree

Redraw with branches as Curve Diagonal Zig



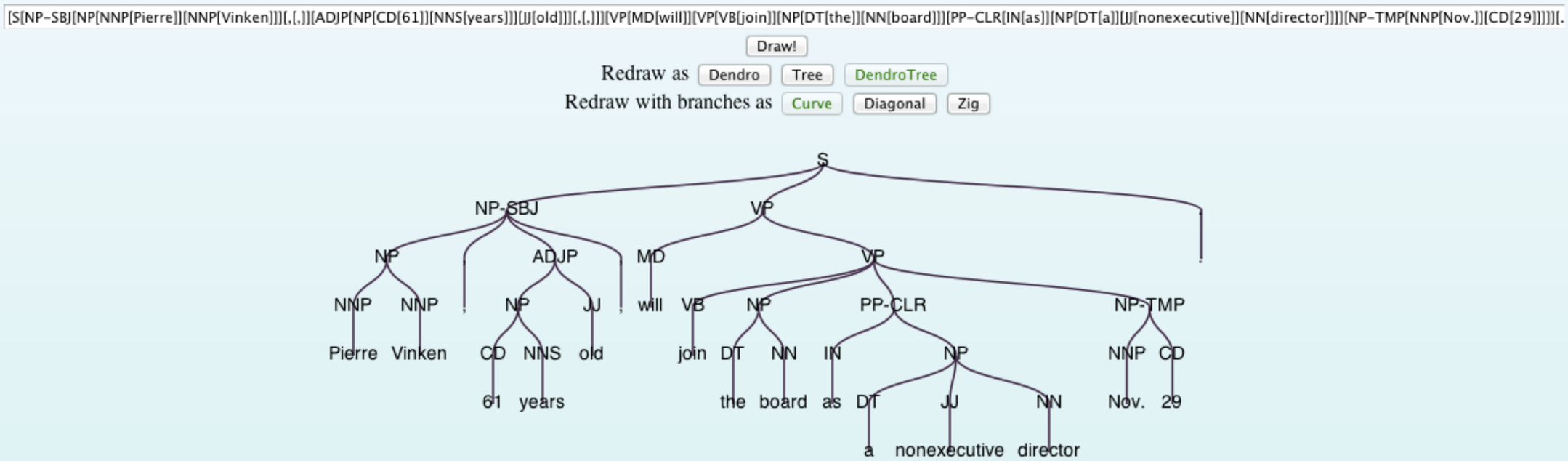
Uses: ProD

Visualizations for tasks and users

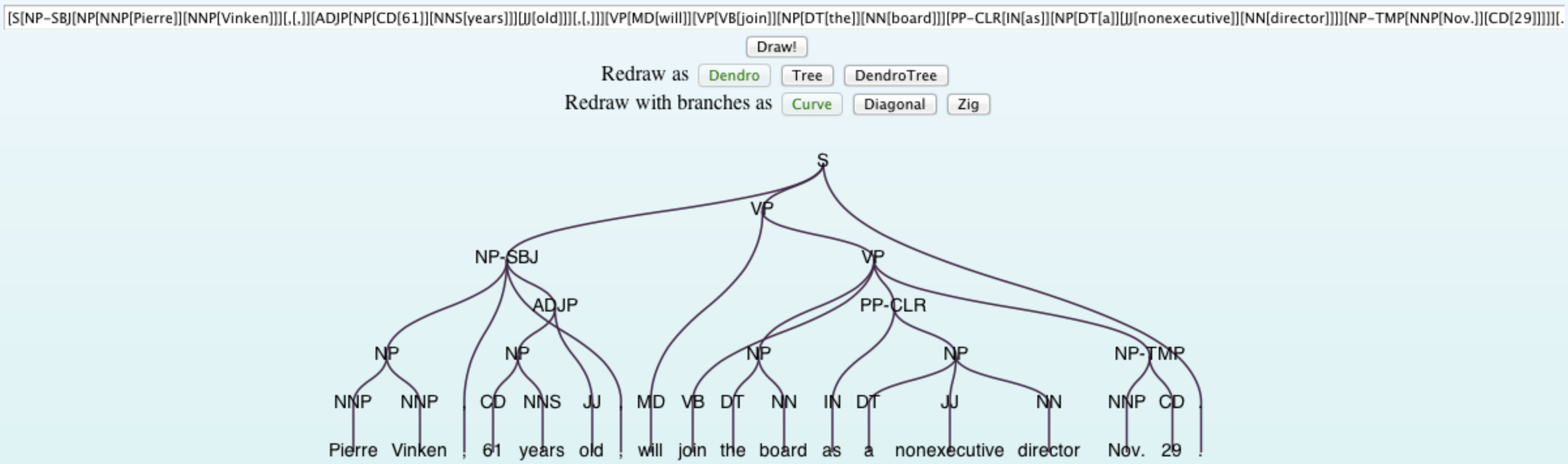


Uses: ProD

Visualizations for tasks and users



Visualizations for tasks and users



Uses: ProD

Visualizations for tasks and users

- Which visualization aspects are primarily and secondarily user preferences?
- **How are conflicts between tasks and user preferences handled?**

Building visualizations

- Sometimes we can use an existing visualization as it is
- Sometimes we can modify an existing visualization
- Sometimes we need to make a new visualization
  - **New visualizations should be reusable components**

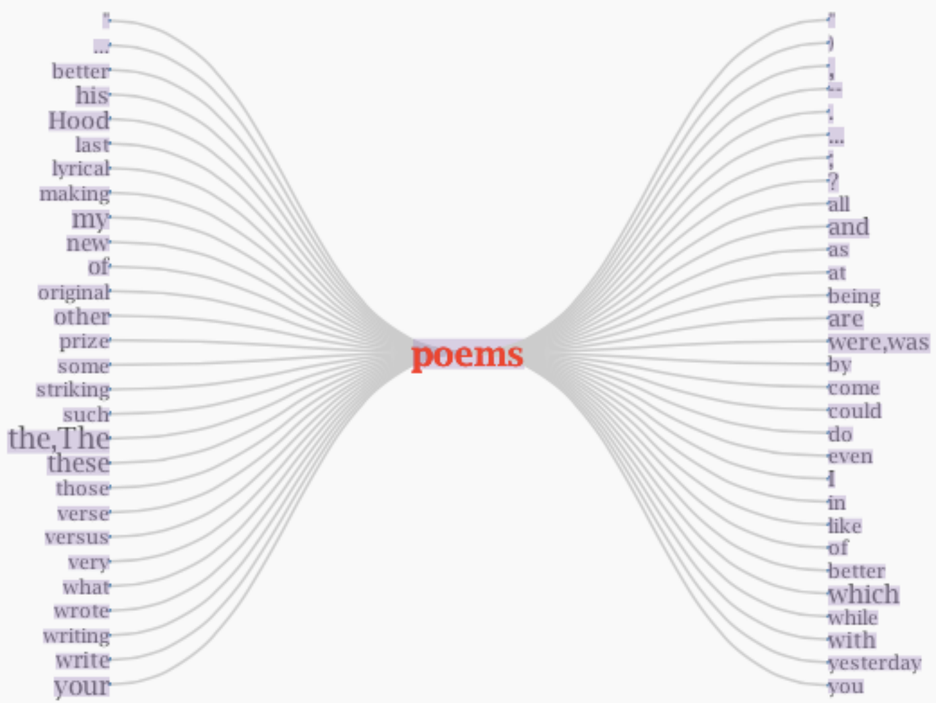
Language visualizations as components

Input file: Choose File no file selected

Word: poem/NNS Find in tree: [aA] Find! Clear

Sort branches by: token

First left cnt ≥: First right cnt ≥: Update filters



Input file: Choose File no file selected

Word: poem/NNS Find in tree: [aA] Find! Clear

Sort branches by: token

First left cnt ≥: First right cnt ≥: Update filters

Uses: DoubleTreeJS

Questions about reusability

- What are the data properties that make a given visualization relevant for the data?
- **What are the fundamental properties and actions of visualizations that form the basis for reusable components?**

i.e. What are the appropriate levels of abstraction to turn a visualization into a component?

References: Grinstein's *WEAVE*, Stasko's *Jigsaw*

Visualizations put ideas into our heads!

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

PHILOSOPHISCHE
FAKULTÄT
Seminar für Sprachwissenschaft

Thank You

christopher.culy@uni-tuebingen.de

http://www.sfs.uni-tuebingen.de/~cculy/