Visualisation of Verb Dependencies

in the EAP Classroom

Yulia Svetashova & Heike da Silva Cardoso

Uni Tübingen

June 21, 2014

1

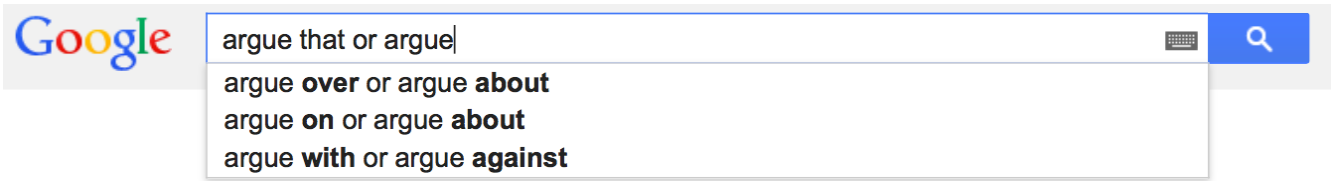
How to Use Concur in a Sentence

By Aya Hadano, eHow Contributor

Like Share 0 Tweet 0 Share Pin it

Attorneys often use the word "concur" in legal documents. It has a long history and its use has evolved over the years. Since "concur" is rarely used in daily conversations, it is important to understand how to correctly use it in a sentence. According to the dictionary, the most common definition of "concur" is "to accord in opinion" or "agree." The most proper way to use "concur" in a sentence is to replace the word "agree."

2



3

23rd July 2012, 4:26 PM

Fenice
Senior Member

Join Date:
Location:
Native language:
Posts:

Hello to all

After contribute, one can either use the infinitive or a form in ING..
I wonder if someone could explain to me how to choose ? Is there a context, a nuance that can guide the choice ?

For instance between these two sentences, is one better (or different) than the other ?
A- Their comparison will contribute to highlight the various drivers of adaptive capacity

B- Their comparison will contribute to highlighting the various drivers of adaptive capacity

Thanks in advance

Fénice

Visualisation of Verb Dependencies

in the British Academic Written English Corpus (BAWE)

I. Processing chain: tokenisation, POS-tagging, parsing

Select a comparison type:

II. Comparison

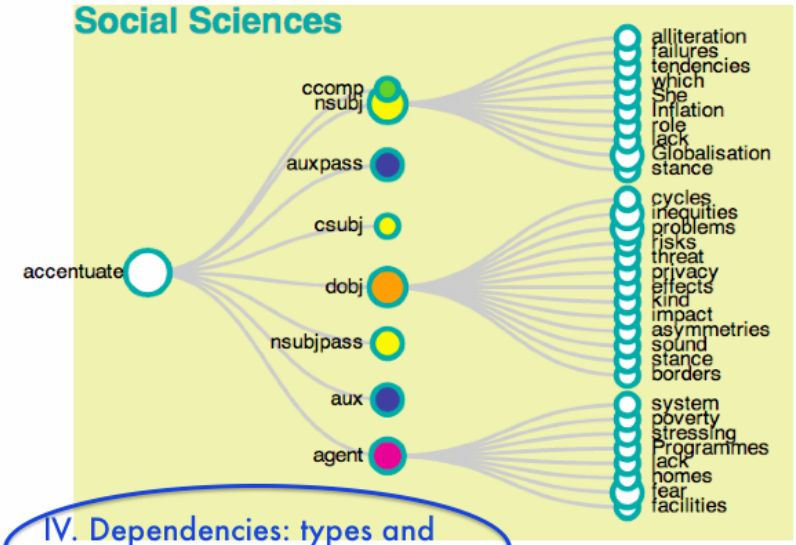
- level of study
- disciplinary group

Select a verb:

- alphabetically
- by frequency

accentuate (59)

III. Frequency data



IV. Dependencies: types and dependents

Contexts

Working class madwomen suffered extreme stigma under the Poor Law in county asylums , whilst the middle and upper class women were more often , when institutionalised , put into private asylums , thus **accentuating** how the ` affluent and middling groups had very different experiences ... -LSB- the -RSB- divide between private and state institutions reflected ... a basic class division ' .23 Ultimately , it is apparent that both perceived causes and treatments of madness differed between the working and middle classes , affecting the lives of many Victorian women .

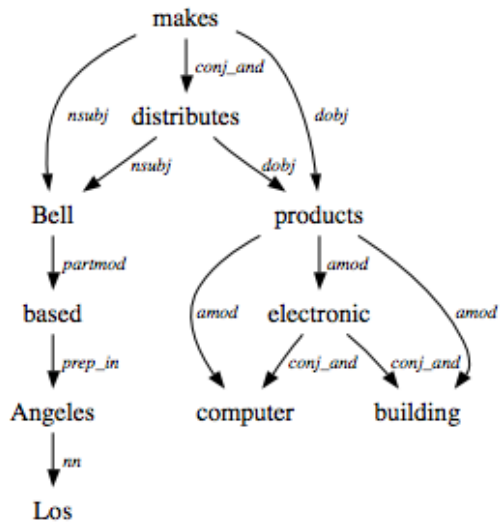
prev next

1 2 3 4 5 6 7

V. Concordance

What is a Dependency Grammar?

In a dependency grammar the verb is seen as the governor or head of a sentence.



Bell, based in Los Angeles, makes and distributes electronic, computer and building products[De Marneffe and Manning(2008)]

Verb Dependencies

Every verb needs a certain number and certain types of arguments, for example:
to eat 2 arguments of which one is optional

- Paul is eating dinner.
- Paul is eating.

while to devour needs two arguments:

- Paul devoured his steak.

and to give needs 3:

- Paul gave Mary the book.

Importance of Arguments

Arguments are essential to the understanding of a verb, because

1. they show us how it can be used and in which context and
2. how the meaning of a verb changes when different arguments are used
3. word sense disambiguation

Why are Verb Dependencies Important in the EAP Classroom?

- How a verb is used helps us understand it better
- If we see how verb usage changes from level 1 to level 4 it helps us deduce what students have to learn and what we can expect of a writer
- By comparing verb usage between different disciplines we can see variations

Stanford Parser

The Stanford Parser [De Marneffe and Manning(2008)] is an algorithm, designed at Stanford University, CA

It takes a sentence as input and gives back a sentence which is annotated with Part of Speech tags [1]

and the dependency structure of a sentence.

[1] POS tags denote the grammatical category of a word e.g. eat = verb, Paul = noun, etc.

Dependency Types

There are different types of dependency relations to a verb, like subject, direct object, indirect object etc.

The Stanford parser assigns these dependency relations to all words in the sentence.

Like in the following sentence:

'She gave me a book'

iobj[2] (gave, me) dobj[3] (gave, book) nsubj[4] (gave, She)

[2] indirect object
[3] direct object
[4] nominal subject

Stanford Dependency Types

The Stanford Dependency Parser uses a set of approximately 50 grammatical relations.

- These relations are strictly binary.
- They hold between a *governor* and a *dependent*.

We chose a subset dependency types, which denote relations between verbs and their arguments

Corpus

We used the BAWE corpus [Nesi and Gardner(2012)] of EAP and ran the Stanford Parser for Dependencies on it.

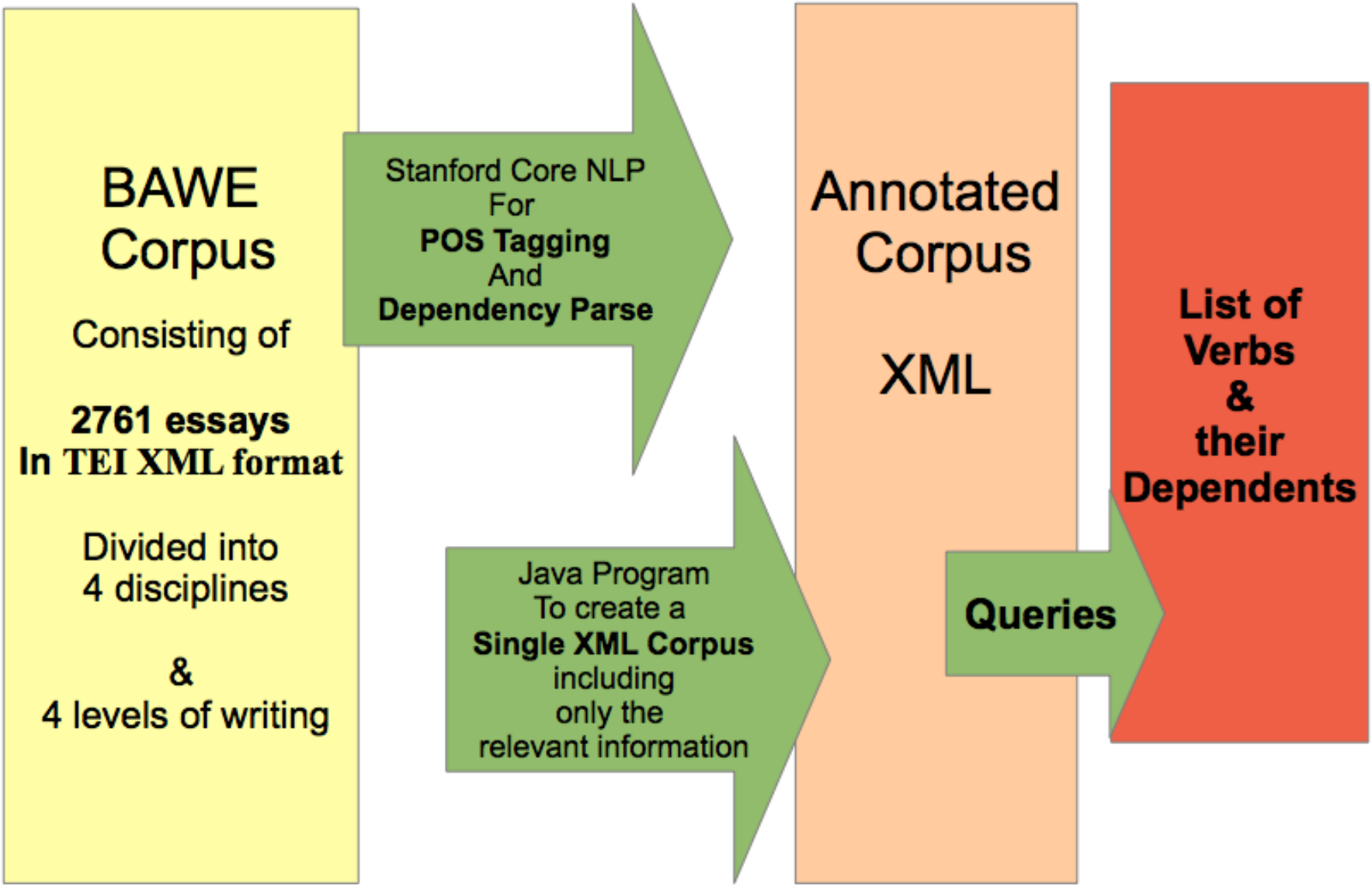
This gave us an annotated corpus including POS tags and dependency relations. This annotated corpus was used for

queries to be able to filter verbs and their dependents for different disciplines and levels of writing.

BAWE Corpus

disciplinary group		Yr 1	Yr 2	Yr 3	Masters	Total
Arts and Humanities	students	101	83	61	23	268
	assignments	239	228	160	78	705
	texts	259	231	161	83	734
	words	468,353	583,617	427,942	234,206	1,714,118
Life Sciences	students	74	71	42	46	233
	assignments	180	193	113	197	683
	texts	191	208	119	203	721
	words	299,370	408,070	263,668	441,283	1,412,391
Physical Sciences	students	73	60	56	36	225
	assignments	181	149	156	110	596
	texts	186	156	169	129	640
	words	300,989	314,331	426,431	339,605	1,381,356
Social Sciences	students	85	88	75	62	313[1]
	assignments	207	197	162	202	777[2]
	texts	218	202	169	204	802
	words	371,473	475,668	440,674	688,921	1,999,130[4]
Total students		**333**	**302**	**234**	**167**	**1039**[1]
Total assignments		**807**	**767**	**591**	**6587**	**2761**[2]
Total texts		**854**	**797**	**618**	**619**	**2897**[3]
Total words		**1,440,185**	**1,781,686**	**1,558,715**	**1,704,015**	**6,506,995**[4]

Processing


BAWE Corpus
Consisting of
2761 essays
In TEI XML format
Divided into 4 disciplines
&
4 levels of writing
Stanford Core NLP For POS Tagging And Dependency Parse
Java Program To create a Single XML Corpus including only the relevant information
Annotated Corpus
XML
Queries
List of Verbs & their Dependents

Why Visualise

1. to visualise patterns we already know exist
2. to visualise structures and detect patterns

Visualisation of Verb Dependency Structures

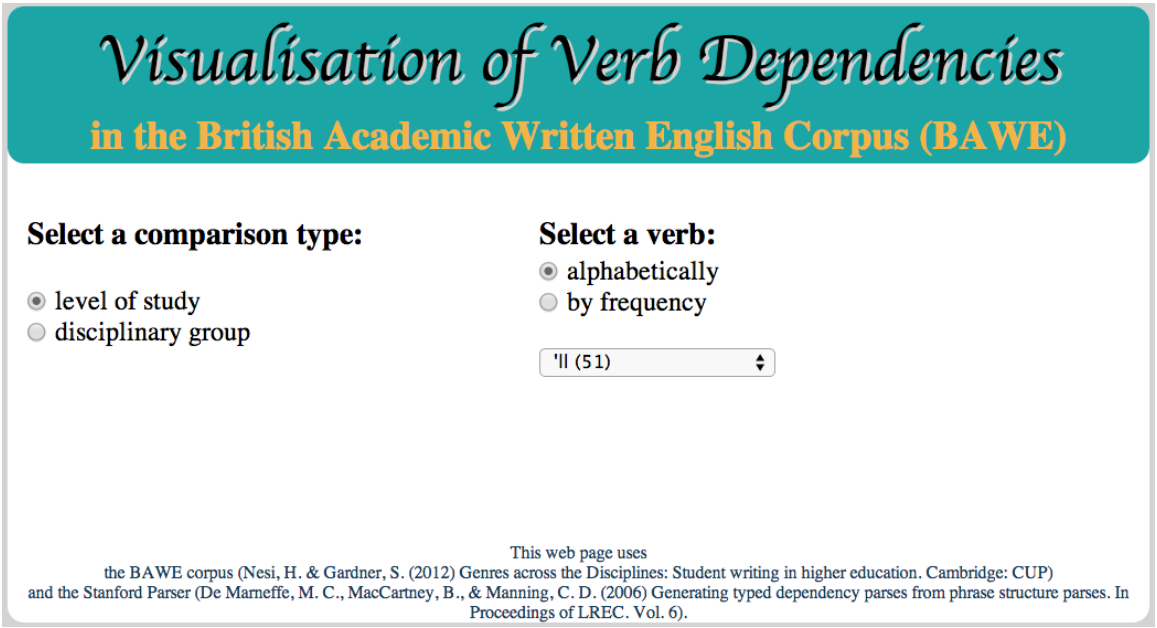
This is a comparative visualisation

1. it shows differences in verb usage
2. it helps to identify specific patterns

Comparison Types

The user can choose between

- level of study comparison: 1, 2, 3, 4 (UG, PG)
- disciplinary groups comparison: Arts and Humanities, Social Sciences, Life Sciences and Medicine, Physical Sciences

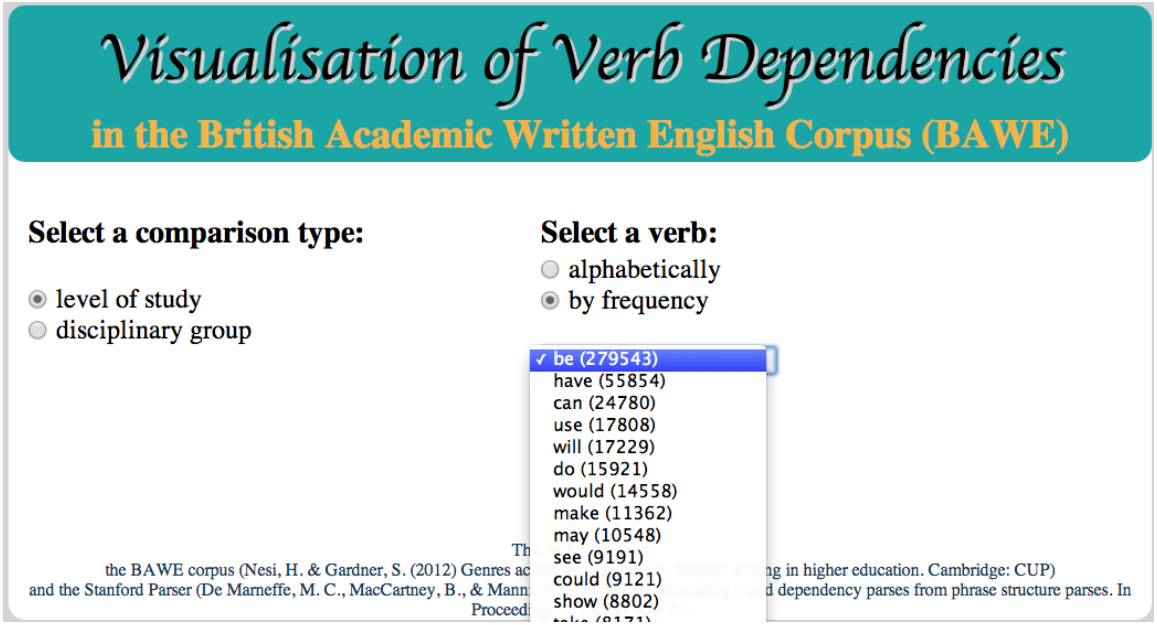


Screenshot of the Initial Page

Lists of Verbs

The user is provided with a list of all verbs in the corpus sorted

- by frequency
- alphabetically



Verb List Sorted by Frequency

Corpus Statistics: Disciplinary Group

Group	Number of Tokens
Arts and Humanities	288,503
Life Sciences	238,205
Social Sciences	337,341
Physical Sciences	235,215
Total (tagged as) verbs:	1,099,264
Total number of words:	6,506,995

- $\mu = 274,816$
- $\approx 17\%$

Corpus Statistics: Level of Study

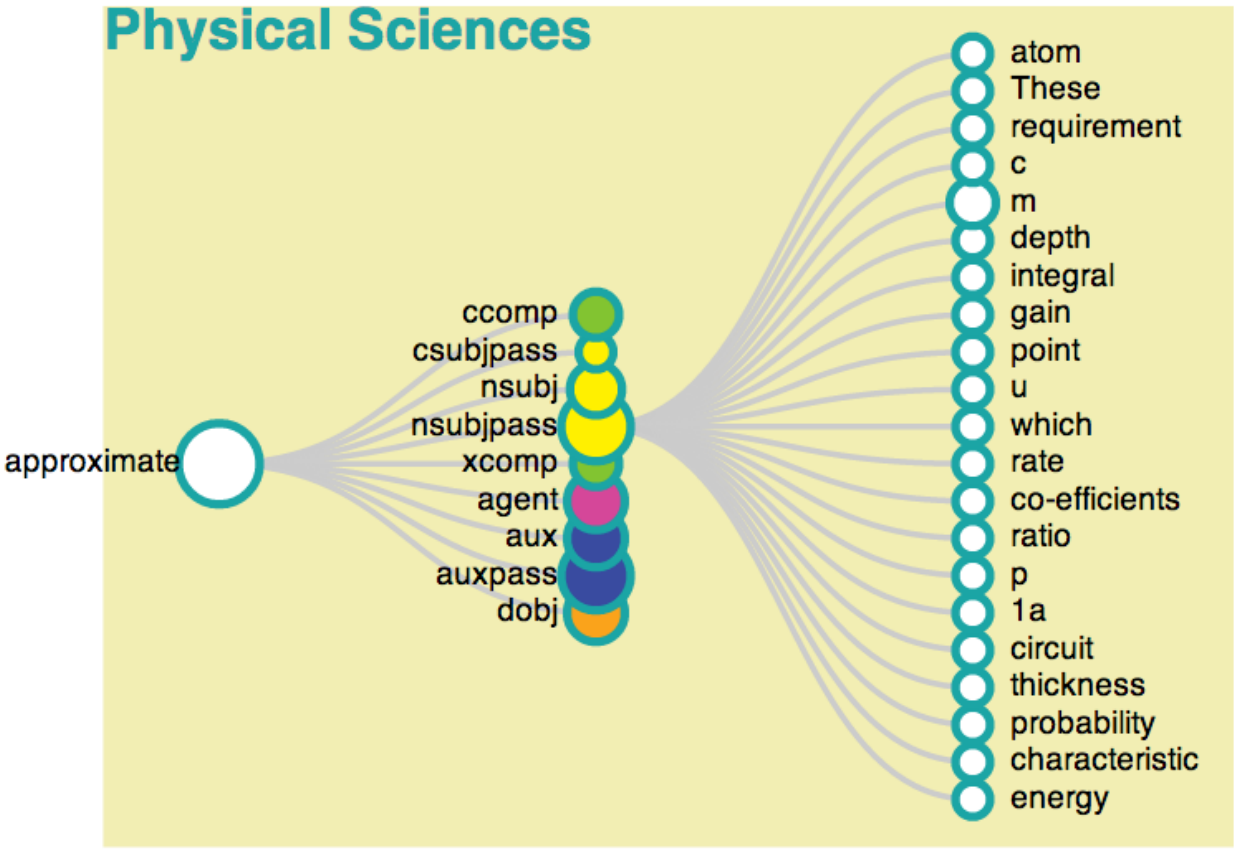
Level of Study	Number of Tokens
1	252,560
2	308,791
3	266,834
4	267,364
unknown	*3,715*
Total (tagged as) verbs:	1,099,264
Total number of words:	6,506,995

- $\mu = 273,887.25$
- $\approx 17\%$

Information Encoded in the Dependency Tree

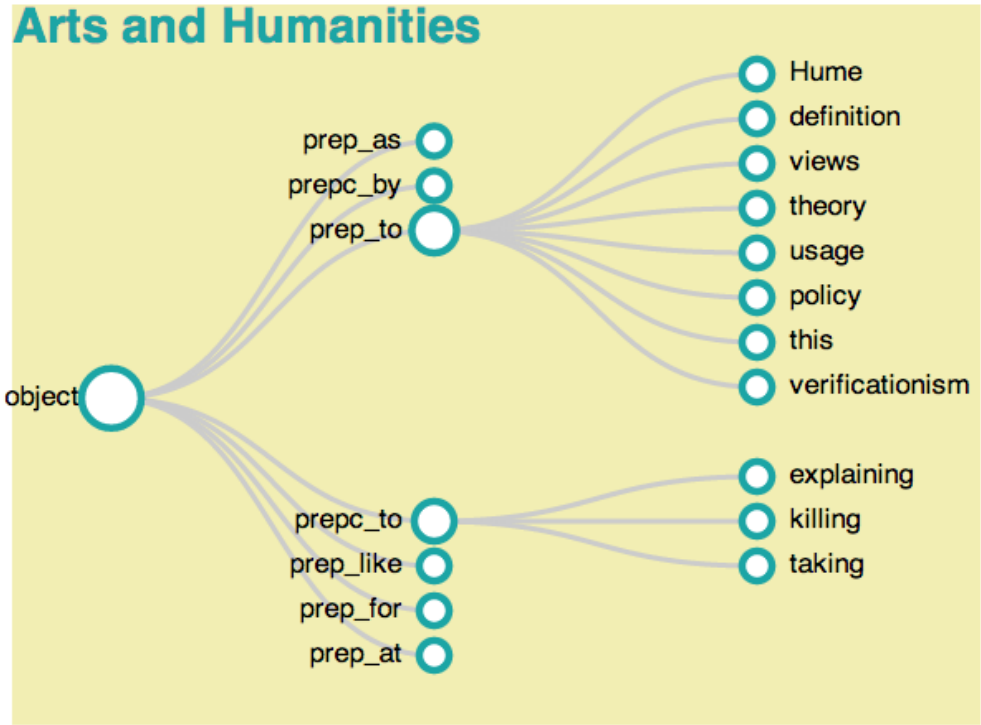
- Verb lemma, its frequency, POS tags, and forms
  - (agree, level 1) *frequency*: 256;
  - *POS tags*: VB, VBN, VBP, VBG, VBD, VBZ;
  - *tokens*: Agreed, agree, agreed, agreeing, agrees
- Dependency relation types, their frequencies
  - (agree, level 1) *type*: nominal subject; *frequency*: 198;
  - *type*: clausal complement; *frequency*: 95, etc.
- Dependents' tokens, POS tags, lemmas, their frequencies
- Contexts of usage

Verb Dependencies Tree



Verb Dependents: Collapsible Tree and KWIC

Verb Dependencies Tree



Contexts

For example , they **object** that if the person is not identical to the human animal , then who is sitting in a chair typing these words , the animal or the person ?

Tim Crane **objected** to Putnam 's Twin Earth theory .

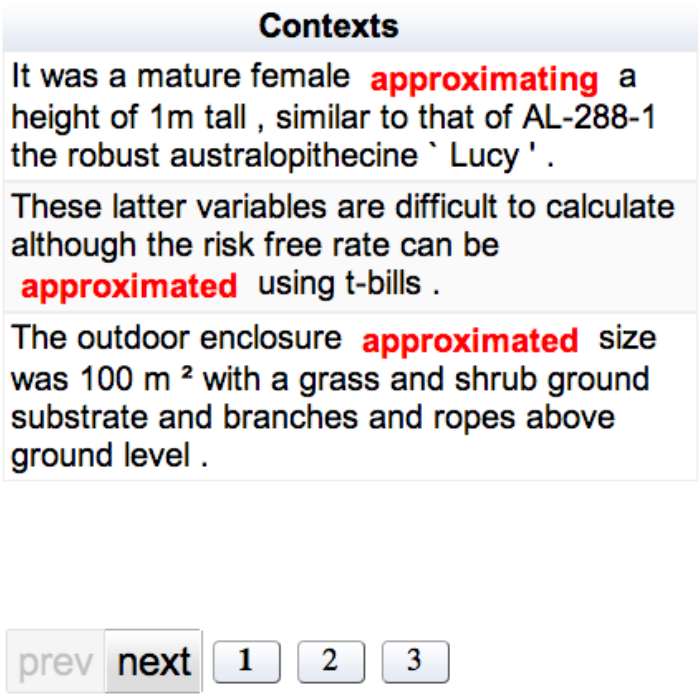
It may be **objected** by proponents of the wager argument that the existence of a made up God is far less probable than the God of Christian theology , but as long as the rewards of worshipping this made up God are infinite , then the value of its outcome is infinite , the same value which worshipping the Christian God carries ; If we were to

prev next

1 2 3 4 5 6 7 8

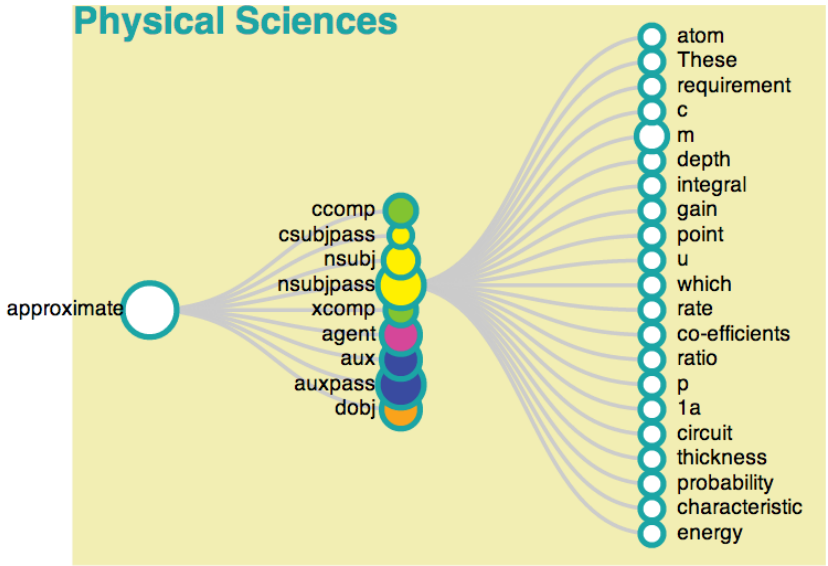
Prepositional Modifiers: Tree and KWIC

Design Decisions



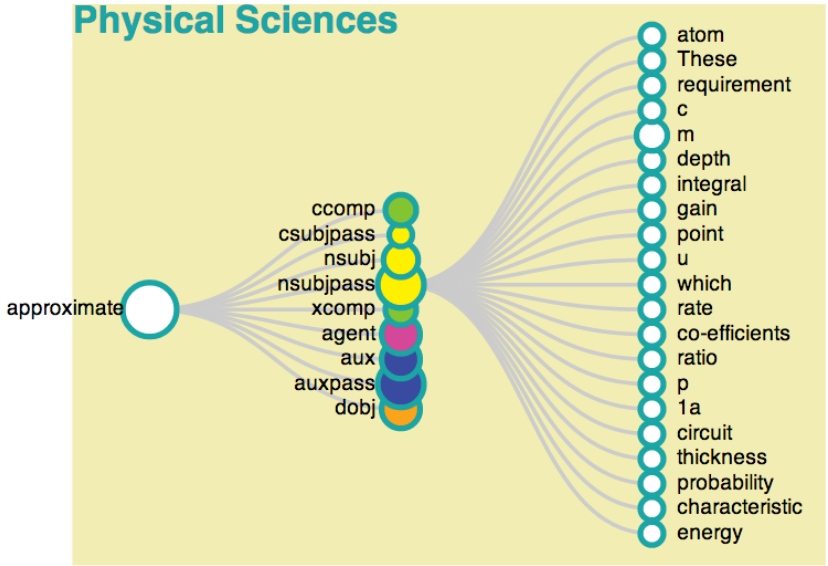
- Visual variable: **color**
  - **selective** interpretation task:
    verbs in the KWIC part are colored red to cause ”pop-up” effect

Design Decisions



- Visual variable: **color**
  - **associative**
    interpretation task:
    the dependency relation types are grouped by meaning (e.g., the group of object tags circles are colored orange, subject group, - yellow, etc.)

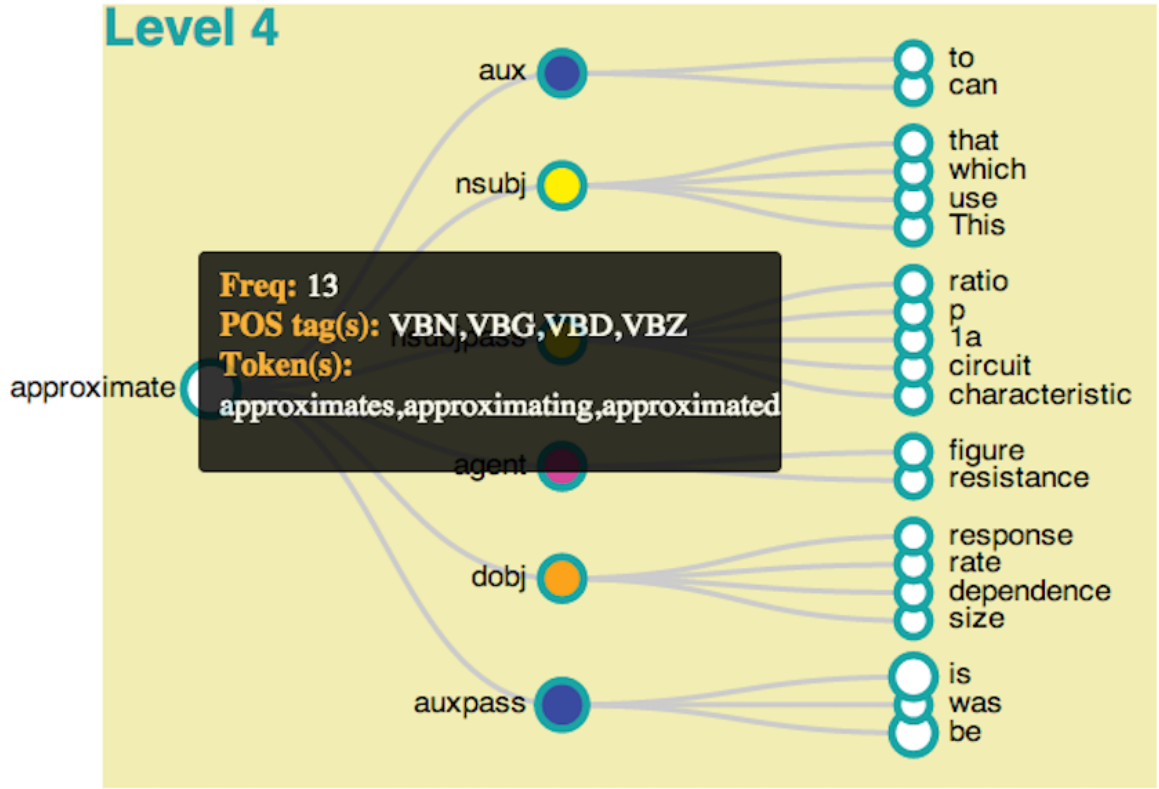
Design Decisions



- Visual variable: **size**
  - **selective** visual interpretation task: elements with higher frequencies have a bigger circle radius

Verb Dependencies Tree

- Hovering the node circle → additional information
- Levels of tree hierarchy: verb lemma, relation type, dependent

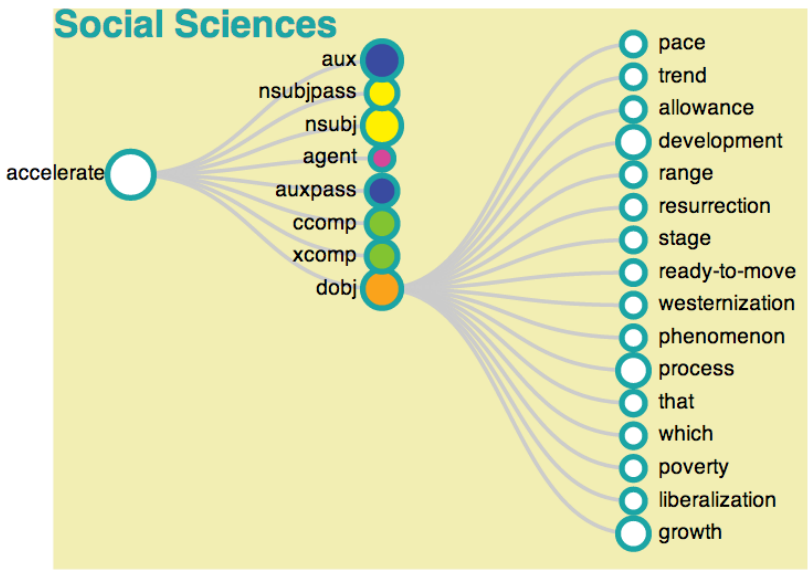


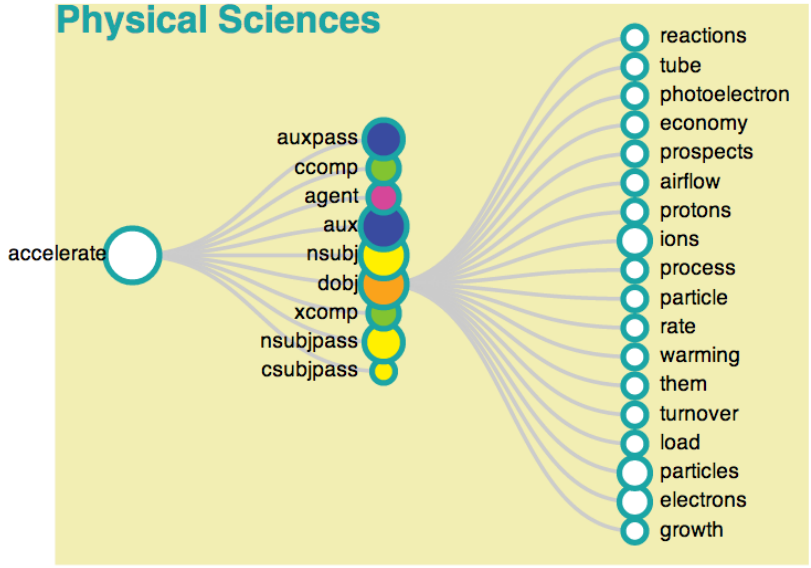
Additional Information

Interaction Types (Yi et al., 2007)

- Select: the verb is selected from a verb list
- Explore: the dependencies information is returned in a form of a tree
- Reconfigure: nodes are collapsible, the user can focus on the most interesting piece of data
- Abstract (Elaborate): data is shown in the tooltips
- Filter: filtering by frequency and alphabetically is possible

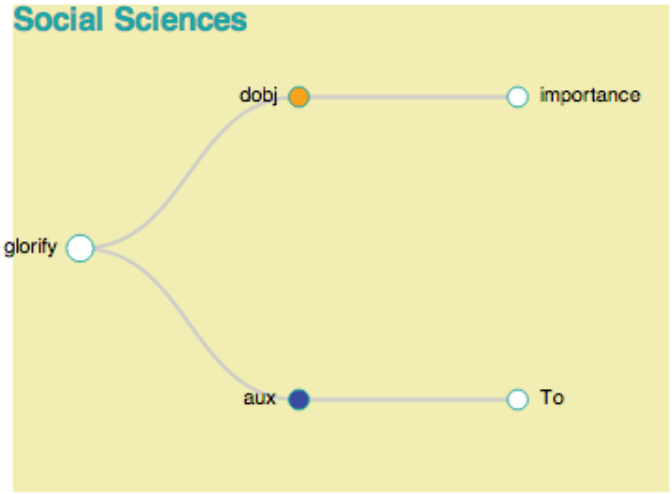
Comparison: Disciplinary Group

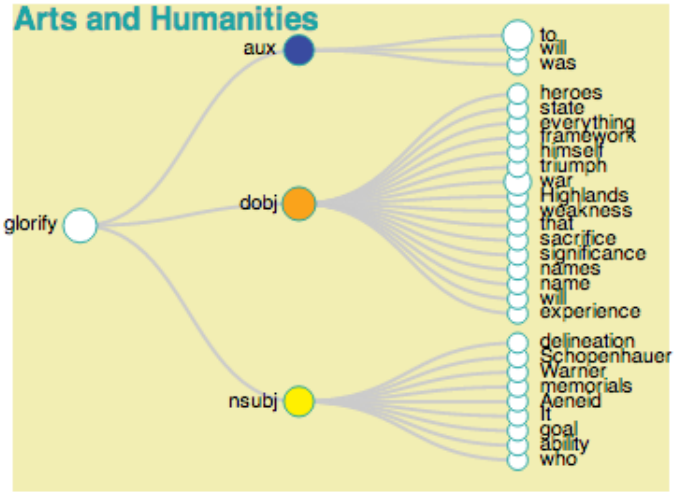




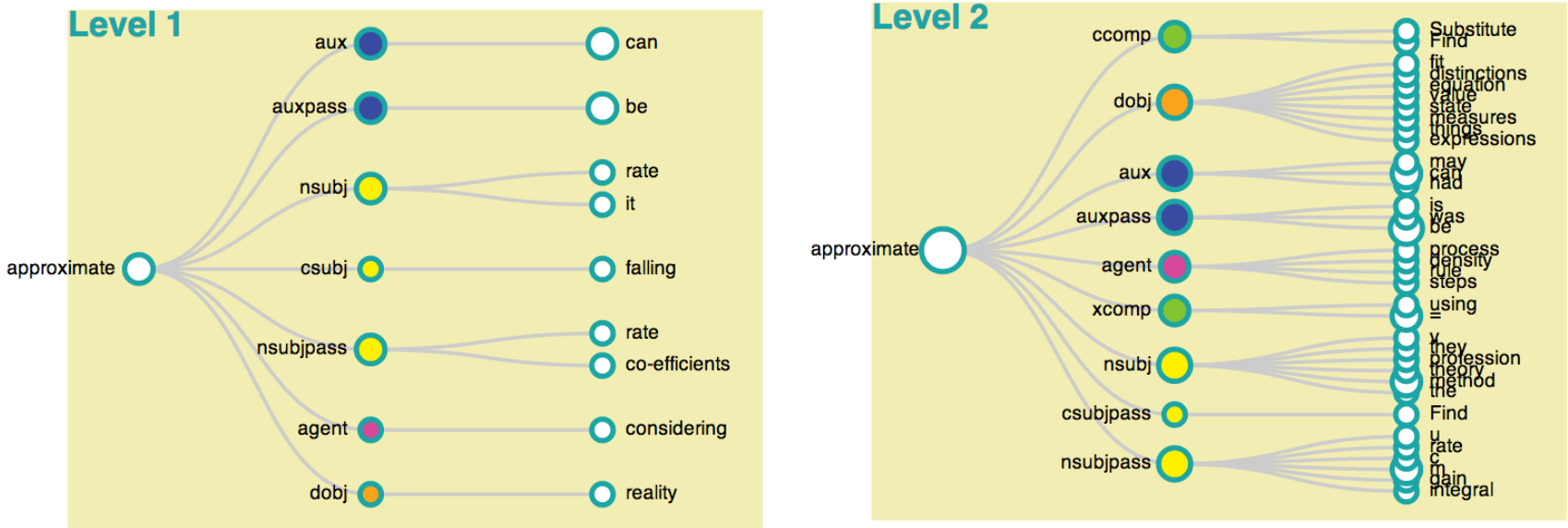
Dependencies Comparison

Comparison: Disciplinary Group





Comparison: Level



Dependencies Comparison

Usage

Two example usages would be to identify

1. styles of writing
2. error patterns

The different types of comparison help identify these features in different fields.

Usage in the EAP Classroom

1. by teachers to help understand the level of writing of their students
2. to see what can be expected in a specific discipline or a level of writing
3. detect frequent errors
4. by students to understand possible word usages

Further Work

- choice options
- more flexible data manipulation
- zooming/scaling
- query data export
- multiple verbs comparison
- adding features to be encoded by visual variables
- pipeline tool
- new levels in tree hierarchy (POS)

Summary

We think that

- Visualisation can be very useful to help identify patterns
- Verb dependency structures capture an essential part of the sentence structure

→ Therefore we believe that the visualisation of verb dependency structures is of interest in many fields


Aufmerksamkeit
attention
Thank
you
for
danke
Ihre
für
your
Merci
mawazo
fyrir
ederiz
virsu
votre
athygli
atención
Siyabonga
yako
jou
uw
Gracias
aandag
Mèsi
þína
kwa
por
teşekkür
atenção
Asante
Dankie
sua
þér
pou
wakho
din
atansyon
Tak
þakka
ilginiz
opmærksomhed
ou
ukunakwa
aandacht
pour
voor
Dank
için
Obrigada

References

MST Carpendale.
Considering visual variables as a basis for information visualisation.
Computer Science TR# 2001-693, 16, 2003.

Marie-Catherine De Marneffe and Christopher D Manning.
Stanford typed dependencies manual.
URL http://nlp. stanford. edu/software/dependencies manual. pdf, 2008.

Hilary Nesi and Sheena Gardner.
Genres across the disciplines: Student writing in higher education.
Cambridge University Press, 2012.

Ji Soo Yi, Youn ah Kang, John T Stasko, and Julie A Jacko.
Toward a deeper understanding of the role of interaction in information visualization.
Visualization and Computer Graphics, IEEE Transactions on, 13(6):1224–1231, 2007.