Computational Grounding of Texts in Real World Proxies

Jason Baldridge
Computational Linguistics Lab
Department of Linguistics
UT Austin
What does “barbecue” mean?
What does “barbecue” *mean*?
What does “barbecue” mean?
What does “barbecue” mean?
What does “barbecue” mean?

Barbecue
What does “barbecue” mean?

Barbecue

- PORK
- BBQ
- HEAT
- GRILLING
- MEAT
- CHARCOAL

COOKING METHODS:
- Wood-fired
- Gas-fired
- Electric

SOUTH AMERICAN ROOTS:
- Slow cooking techniques

WORLDWIDE POPULARITY:
- Number of barbecue events (in billions)
  - 1987: 1.4
  - 1991: 2.3
  - 1993: 2.6
  - 1995: 2.7
  - 1997: 2.9

© 2011 Jason M Baldridge

Friday, May 27, 2011
What does “barbecue” mean?

Barbecue'
What does “barbecue” mean?

Barbecue’

COOKING
BBQ
WOOD
GRILLING
SMOKING
CHARCOAL
MEAT
GAS
WATER
PROPANE
SOUTH
USUAL
BARBECUES
AMERICAN
BARBECUE
STYLED
SMOKED
COOKED
BRAISING
APPLIED
FLAVORS
FRENCH
MEATS

NUMBER OF BARBECUE EVENTS (in Billions)

<table>
<thead>
<tr>
<th>Year</th>
<th>1987</th>
<th>1991</th>
<th>1993</th>
<th>1995</th>
<th>1997</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>1.4</td>
<td>2.3</td>
<td>2.6</td>
<td>2.7</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Friday, May 27, 2011
Travel at the Turn of the 20th Century

http://www.lib.utexas.edu/books/travel/index.html
Grapes of Wrath in Google Earth

Grapes of Wrath in Google Earth

A SONG of the rolling earth, and of words according,
Were you thinking that those were the words, those upright lines? those curves, angles, dots?
No, those are not the words, the substantial words are in the ground and sea,
They are in the air, they are in you.
Annotation of Memoranda During the War

DURING the Union War I commenced at the close of 1862, and continued steadily through ’63, ’64 and ’65, to visit the sick and wounded of the Army, both on the field and in the Hospitals in and around Washington city. From the first I kept little note-books for impromptu jottings in pencil to refresh my memory of names and circumstances, and what was specially wanted, &c. In these I brief’d cases, persons, sights, occurrences in camp, by the bedside, and not seldom by the corpses of the dead. Of the present Volume most of its pages are renderings from such pencillings on the spot. Some were scratch’d down from narratives I heard and itemized while watching, or waiting, or tending somebody amid those scenes. I have perhaps forty such little note-books left, forming a special history of those wars, for myself alone.
<table>
<thead>
<tr>
<th>Time</th>
<th>Persons</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00:54</td>
<td>Yves</td>
<td>Could you please tell us your name?</td>
</tr>
<tr>
<td>00:00:55</td>
<td></td>
<td>Je m’appelle [redacted]. Avant la guerre... je veux dire durant la guerre, je vivais ici dans cette cellule. Avant elle était appelée Gasharu Il mais aujourd’hui elle s’appelle Kagina, dans le secteur de Kicukiro, et le district Kicukiro. J’habitais chez mon oncle et lorsque la guerre a commencé, nous avons fui de la maison où nous habitons et sommes allés chercher refuge dans une autre maison voisine. Les Interahamwe sont venus après nous. Mon oncle a été tué dans les attaques. Nous avons été chassés et nous avons dû aller à la vie. Tous les Interahamwe sont venus après nous. Ils nous ont menacés et nous ont menacés. Ils étaient plus pressés que nous et je me suis glissé en se pousser pour nous pas loin de Interahamwe et ils ont arrêté le jour suivant, je suis retourné à la maison.</td>
</tr>
<tr>
<td>00:01:36</td>
<td></td>
<td>Il nous ont emmenés à la maison. Ils étaient plus pressés que nous et je me suis glissé en se pousser pour nous pas loin de Interahamwe et ils ont arrêté le jour suivant, je suis retourné à la maison.</td>
</tr>
<tr>
<td>00:02:11</td>
<td>Yves</td>
<td>Did you say that seven people?</td>
</tr>
<tr>
<td>00:02:15</td>
<td></td>
<td>More than seven people</td>
</tr>
<tr>
<td>00:02:15</td>
<td></td>
<td>Plus de 7 personnes</td>
</tr>
</tbody>
</table>

**Amagambo**

<table>
<thead>
<tr>
<th>Ubaza</th>
<th>Ubaza/Ubazwa</th>
<th>Amagambo</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00:54</td>
<td>Yves</td>
<td>Ushobora kutubwira amazina anyyu?</td>
</tr>
<tr>
<td>00:00:55</td>
<td>[redacted]</td>
<td>Nyewe nitwa kaba mbere y’intambara mu ntambara nari rnye hano mu kagali, mbere kari Gasharu ya kabiri arko ubu ni Kagina, umurenge wa Kicukiro, akarete ka Kicukiro. Nabadaga kwa marume noneho intambara ibaye, turahungu tuva mu nzu twagambo turahungu tuva mu yindi ihi yayo, noneho interahamwe iriza zihi patrihura ziratuzana marume bwicicira kuri etaje ihiyoyi hano.</td>
</tr>
<tr>
<td>00:01:36</td>
<td>[redacted]</td>
<td>Baratuzana abandri twari kumiwe babicira hano ndumva bararengara barindwa. Bamaze kubica ubwo nje nahe mbcakwa, nyuha hagati y’amazu hira aho ngera ku mubanda wa hano hepfo wigataka Ubwo interahamwe zirengare nk’ici, zahise zimbahoza zimaze kumbohoza zirandara buracya naignenda nsibira hepfo y’ahoro twagambo.</td>
</tr>
<tr>
<td>00:02:11</td>
<td>Yves</td>
<td>Ahangaha mwasubwoni ko haye abantu barenge barindwi.</td>
</tr>
<tr>
<td>00:02:15</td>
<td>[redacted]</td>
<td>Haguye abantu barenge barindwi.</td>
</tr>
</tbody>
</table>
Hôtel des Mille Collines

The Hôtel des Mille Collines is a large hotel in Kigali, the capital city of Rwanda. It became famous as the building in which more than a thousand people took refuge during the Rwandan Genocide of 1994. The story of the hotel and its manager at that time, Paul Rusesabagina, was used as the basis of the film Hotel Rwanda.

Full Article

All text is available under the terms of the GNU Free Documentation License.
TextGrounder: location prominence

...thermometer stood at 130 at dusk on the river. It was just pleasantly cool. We woke next morning at the San Francisco wharf where the cold sea-breeze made overcoats a necessity. The seasons are all mixed up in that city in August...
## Automatically extracted topics from 20th Century Travel

<table>
<thead>
<tr>
<th>WATER</th>
<th>LAND</th>
<th>WRITING</th>
<th>ART</th>
<th>EGYPT</th>
<th>CITY</th>
<th>MECCA</th>
<th>PLAINS</th>
</tr>
</thead>
<tbody>
<tr>
<td>sea</td>
<td>mountain</td>
<td>cloth</td>
<td>church</td>
<td>egypt</td>
<td>street</td>
<td>mecca</td>
<td>city</td>
</tr>
<tr>
<td>coast</td>
<td>feet</td>
<td>book</td>
<td>renaissance</td>
<td>nile</td>
<td>london</td>
<td>city</td>
<td>sioux</td>
</tr>
<tr>
<td>island</td>
<td>rock</td>
<td>illustrations</td>
<td>madonna</td>
<td>cairo</td>
<td>station</td>
<td>place</td>
<td>chicago</td>
</tr>
<tr>
<td>steamer</td>
<td>valley</td>
<td>story</td>
<td>picture</td>
<td>egyptian</td>
<td>railway</td>
<td>day</td>
<td>river</td>
</tr>
<tr>
<td>bay</td>
<td>great</td>
<td>books</td>
<td>venice</td>
<td>temple</td>
<td>road</td>
<td>place</td>
<td>texas</td>
</tr>
<tr>
<td>land</td>
<td>mountains</td>
<td>work</td>
<td>mark</td>
<td>ancient</td>
<td>time</td>
<td>time</td>
<td>kansas</td>
</tr>
<tr>
<td>ship</td>
<td>hills</td>
<td>life</td>
<td>doge</td>
<td>time</td>
<td>hotel</td>
<td>desert</td>
<td>missouri</td>
</tr>
<tr>
<td>harbour</td>
<td>rocks</td>
<td>history</td>
<td>early</td>
<td>rameses</td>
<td>square</td>
<td>city</td>
<td>west</td>
</tr>
<tr>
<td>port</td>
<td>side</td>
<td>author</td>
<td>saints</td>
<td>water</td>
<td>place</td>
<td>mahomet</td>
<td>business</td>
</tr>
<tr>
<td>shore</td>
<td>miles</td>
<td>crown</td>
<td>century</td>
<td>desert</td>
<td>house</td>
<td>country</td>
<td>miles</td>
</tr>
<tr>
<td>boat</td>
<td>plain</td>
<td>illustrated</td>
<td>great</td>
<td>tombs</td>
<td>house</td>
<td>arabia</td>
<td>hundred</td>
</tr>
<tr>
<td>water</td>
<td>river</td>
<td>gilt</td>
<td>work</td>
<td>years</td>
<td>great</td>
<td>journey</td>
<td>dakota</td>
</tr>
<tr>
<td>board</td>
<td>hill</td>
<td>edition</td>
<td>wall</td>
<td>english</td>
<td>office</td>
<td>days</td>
<td>streets</td>
</tr>
<tr>
<td>islands</td>
<td>stone</td>
<td>vols</td>
<td>works</td>
<td>temples</td>
<td>regent</td>
<td>pilgrims</td>
<td>railroad</td>
</tr>
<tr>
<td>town</td>
<td>long</td>
<td>volumee</td>
<td>room</td>
<td>luxor</td>
<td>club</td>
<td>camels</td>
<td>lincoln</td>
</tr>
</tbody>
</table>
TextGrounder: Location-specific topics [Speriosu et al. 2010]
Geographic meaning of words (based on Wikipedia)
Geographic meaning of words (based on Wikipedia)
Geographic meaning of words (based on Wikipedia)
Geographic meaning of words (based on Wikipedia)
Geographic meaning of words (based on Wikipedia)

beach

mountain
Geographic meaning of words (based on Wikipedia)

mountain

beach

mountain

beach
Geographic meaning of words (based on Wikipedia)

- Beach
- Mountain
- Wine
Geographic meaning of words (based on Wikipedia)

- beach
- mountain
- wine
Geographic meaning of words (based on Wikipedia)

- beach
- mountain
- wine
- barbecue
Regional variations in words on Twitter
Automatic geolocation of Wikipedia pages and Twitter users
Wing and Baldridge 2011, ACL
Automatic geolocation of Wikipedia pages and Twitter users
Wing and Baldridge 2011, ACL
Temporal resolution of Wikipedia biographies (in submission)
Temporal resolution of Wikipedia biographies (in submission)
Temporal resolution of Wikipedia biographies (in submission)
Temporal resolution of Wikipedia biographies (in submission)
Temporal resolution of Wikipedia biographies (in submission)

Abraham Lincoln

This article is about the American president. For other uses, see Abraham Lincoln (disambiguation).

Abraham Lincoln was a politician who served as the 16th President of the United States. He was one of the most influential figures in American history, and he has been widely studied and debated for his role in the Civil War and his efforts to end slavery. Lincoln was known for his strong leadership and his ability to unite the nation during a time of great crisis. He is remembered for his famous Gettysburg Address, in which he articulated the Union's commitment to the ideals of liberty and equality.

Chronons

P(x|d)

2000 AD

4000 BC

2000 BC

0 AD

1835 AD

0.005291775

1700 AD 1715 AD 1730 AD 1745 AD 1760 AD 1775 AD 1790 AD 1820 AD 1835 AD 1850 AD 1865 AD 1870 AD 1880 AD 1885 AD 1890 AD 1895 AD 1900 AD 1905 AD 1910 AD 1915 AD 1920 AD 1925 AD 1930 AD 1935 AD 1940 AD 1945 AD 1950 AD 1955 AD 1960 AD 1965 AD 1970 AD 1975 AD 1980 AD 1985 AD 2000 AD
Temporal resolution of Wikipedia biographies (in submission)
Temporality of words, by hour http://www.tweetolife.com/hour/
Temporality of expressions, by day: http://www.google.com/trends
Temporality of expressions, by day: http://www.google.com/trends
Temporality of expressions, by year: http://ngrams.googlelabs.com/

slave  trenches  aircraft  war

© 2011 Jason M Baldridge

TILTS Symposium, May 2011

Friday, May 27, 2011
Lexical brain decoding [Yarkoni, Poldrack, Nichols, Van Essen & Wager (in press)]
Lexical brain decoding [Yarkoni, Poldrack, Nichols, Van Essen & Wager (in press)]

Term-based search

“Pain” →

Related studies

Mechanisms of Directed
An fMRI Investigation of
Placebo-Induced Changes in fMRI
in the Anticipation and
Experience of Pain

Automated coordinate extraction

\begin{tabular}{cccc}
X & Y & Z & Study \\
-23 & 18 & 45 & 1 \\
19 & 3 & 12 & 1 \\
-40 & 0 & -16 & 1 \\
35 & -41 & 29 & 2 \\
2 & 18 & 33 & 2 \\
\end{tabular}

Meta-analysis

\[ P(\text{Pain Activation}) \]

Classification

Working mem. \( P = 78\% \)

Emotion \( P = 64\% \)

Pain \( P = 87\% \)

Select highest probability

→ “Pain”
Scaling up: more documents, more context (more problems?)
Scaling up: more documents, more context (more problems?)
Scaling up: more documents, more context (more problems?)
Scaling up: more documents, more context (more problems?)
Scaling up: more documents, more context (more problems?)
Scaling up: more documents, more context (more problems?)
Scaling up: more documents, more context (more problems?)

THE KEPT

"keep" "delete"

THE DELETED

"keep" "delete"