Engishes of Science: Concepts, Analyses, Teaching Applications

Albanian Society for the Study of English
Vlore 2011

Josef Schmied
Chair English Language & Linguistics
Chemnitz University of Technology
www.tu-chemnitz.de/phil/english/schmied
josef.schmied@phil.tu-chemnitz.de
Survey

1. definitions and framework: Englishes of Science
2. database: the SPACE corpus (=Specialised and Popular ACademic English)
3. central linguistic concepts in academic writing: complexity, stance/hedging/modality, cohesion
4. analysis: differences between specialized and popular text versions in complexity, hedging, cohesion
5. teaching: discovery procedures, guidelines and exercises, MA module
6. conclusion: outlook
1. Definitions and framework

1.1 Social constructivism

Social constructionism and **social constructivism** are **sociological theories** of **knowledge** that consider how social phenomena develop in social contexts. Within constructionist thought, a **social construction** (social construct) is a concept or practice that is the construct (or **artifact**) of a particular group. When we say that something is socially constructed, we are focusing on its dependence on contingent variables of our social selves.

Constructivism applied to Engishes of Science

writers and readers interact in a discourse community on the basis of accepted institutionalised conventions of metadiscourse

- Swales (1990): genre approach to academic writing
- Hyland (2005, 2009): author stance and engagement are crucial variables in academic interaction
  “writing is always a personal and socio-cultural act of identity whereby writers both signal their membership in a range of communities as well as express their own creative presence” (Hyland 2006: 35)
1.2 Discourse community
(Swales 1990: 24-27)

- A discourse community utilizes and hence possesses one or more genres in the communicative furtherance of its aims.
- In addition to owning genres, a discourse community has acquired some specific lexis.
- A discourse community has a threshold level of members with a suitable degree of relevant content and discoursal expertise.

→ specialised and popular academic discourse
1.3 Englishes of Science according to discourse community

- research discourse
- popular discourse
- newspaper science discourse
- student discourse
- instructional discourse

popularisation
1.4 Englishes of Science according to genre/text-types in community

**research “output”**
- research articles
  - book reviews
- books and handbook articles
- project proposals, reports
- conference presentations

**teacher “talk”/e-learning**
- ppt presentations
  - lectures
- student presentations
- textbooks
- www pages
- Wikis and other platforms

**science “journalism”**
- popular science articles
- popular science books
- newspaper science articles

**student "literacy"**
- fieldwork notes/essays
- Mag/BA thesis
- seminar presentations
1.5 Practical examples of discourse-specific science Englishes: 0048PN

PNAS, 2001:
Tumor cell surface heparan sulfate as cryptic promoters or inhibitors of tumor growth and metastasis

NewScientist, 2002:
Pass the sugar

PNAS commentary, 2002:
Six blind men and the elephant—the many faces of heparan sulfate

TimesOnline, 2005:
Cancer drug 'smart cell' can attack tumours from inside

MailOnline, 2005:
The microscopic anti-cancer 'smart bomb'
1.5 Practical examples of discourse-specific science Englishes: 0084PN

PNAS, 2002:
Efficient production by sperm-mediated gene transfer of human decay accelerating factor (hDAF) transgenic pigs for xenotransplantation

Daily Mail, 2002:
Scientists hail transplant breakthrough

BBC, 2002:
Technique offers transplant hope

Nature, 2003:
Piglets add some colour to transgenic story
2. The Corpus of Specialised and Popular ACademic English (SPACE)

**Rationale**: compare expert texts with the “same content” in expert-to-expert and expert-to-academic layperson communication

- science journals like *New Scientist* (subscription)
- academic online databases, pre-publication servers like *arXiv* (arxiv.org)
- publications in the *Proceedings of the National Academy of Sciences* (PNAS, pnas.org)
- plus: e.g. *Public Library of Science - Medicine* (plos.org)
Condition critical: The medical crisis facing America
A nation addicted to medical technology faces a health crisis that could drive it to economic ruin – will the next president have a cure?

North American fish swimming towards extinction
A third of freshwater fish in the region have become more endangered over the last 20 years, as human activity fragments rivers and lakes
10:44 20 September 2008

‘LHC day’ was highest profile physics event in history
A billion people tuned in to watch the Large Hadron Collider start up – now physicists hope the project will inspire a new generation of scientists
09:13 20 September 2008 6 comments

Clearer skies have brought more rain
Air pollution has dropped over recent decades, and the extra sunlight entering the atmosphere has led to a steady rise in average rainfall over land
08:19 20 September 2008 7 comments

Cosmic explosion is most distant ever seen
The most distant gamma-ray burst ever detected exploded 12.8 billion light years from Earth
01:22 20 September 2008 1 comment

Why ecotourists should head to high-rise resorts
Ecotourists might shun high-rise holiday resorts, but tourist towns could be better for the environment than discreet villas
18:00 19 September 2008 3 comments
### Table 1: Domain set-up of the SPACE10 corpus

<table>
<thead>
<tr>
<th>texts</th>
<th>readership</th>
<th>specialised</th>
<th>popular</th>
<th>popular: special</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>words</td>
<td>words</td>
<td>percentage</td>
</tr>
<tr>
<td>physics</td>
<td></td>
<td>AX</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>quantum-</td>
<td>11</td>
<td>115981</td>
<td>15327</td>
<td>13%</td>
</tr>
<tr>
<td>particle-</td>
<td>2</td>
<td>26384</td>
<td>4574</td>
<td>17%</td>
</tr>
<tr>
<td>astro-</td>
<td>9</td>
<td>167553</td>
<td>18119</td>
<td>11%</td>
</tr>
<tr>
<td>bioscience</td>
<td></td>
<td>PN</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>biochemistry</td>
<td>6</td>
<td>53212</td>
<td>7114</td>
<td>13%</td>
</tr>
<tr>
<td>genetics</td>
<td>24</td>
<td>183312</td>
<td>19101</td>
<td>10%</td>
</tr>
<tr>
<td>microbiology</td>
<td>10</td>
<td>53139</td>
<td>5176</td>
<td>10%</td>
</tr>
<tr>
<td>∑</td>
<td>62</td>
<td>599581</td>
<td>69411</td>
<td>12%</td>
</tr>
<tr>
<td>average</td>
<td></td>
<td>99930</td>
<td>11568</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: SPACE07 text publication dates

<table>
<thead>
<tr>
<th>Month/Year</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov 00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec 00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mar 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>May 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jul 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aug 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec 01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mar 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>May 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jul 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aug 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec 02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mar 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>May 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jul 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aug 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec 03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mar 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>May 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jul 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aug 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec 04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan 05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb 05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mar 05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr 05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>May 05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
**0090PN** Topical DNA oligonucleotide therapy reduces UV-induced mutations and photocarcinogenesis in hairless mice

WT repair-proficient JJ and CC heterozygous JJ partially RB repair-deficient JJ mice NNS both DT transgenic NN for IN the DT mutation-indicator NN gene NN The DT ability NN of IN topically RB applied VVN pTT NN to TO induce VV protective JJ physiologic JJ responses NNS that WDT normally RB result VVP from IN DNA NN damage NN suggests VVZ_M a DT previously RB undescribed JJ means NNS of IN reducing VVG skin NN cancer NN in IN high-risk JJ individuals NNS SENT . Skin NN cancer NN accounts NNS for IN at IN least JJS 40 CD % NN of IN all DT human JJ malignancies NNS , 1,000,000 JJ cases NNS annually RB in IN the DT U.S. NP Incidence NN is VBZ clearly RB linked VVN to TO UV NP UV exposure NN and CC increases NNS exponentially RB with IN age NN 1 CD , 3 CD SENT . Skin NN cancer NN risk NN is VBZ greatly RB increased VVN in IN the DT rare JJ disease NN xeroderma NN pigmentosum NN

**0090NS** Suntan lotion primes the skin's defences

IT PP MIGHT MD_3 be VB possible JJ_3 to TO develop VV suntan NN lotions NNS that WDT kick-start VV - the DT skin NN 's POS protective JJ mechanisms NNS against IN cancer NN before IN you PP hit VVD the DT beach NN The DT key JJ ingredient NN could MD_3 be VB a DT fragment NN of IN DNA NP just RB two CD bases NNS long RB called VVD a DT TT NP TT dimer NN that WDT mimics VVZ one CD of IN the DT signs NNS of IN DNA NN damage NN from IN ultraviolet JJ light NN People NNS who WP want VVP a DT tan NN may MD not RB even RB need VVP to TO go VV out RP in IN the DT sun NN . Mouse NN skin NN does VVZ not RB produce VV melanin NN but CC earlier JJR tests NNS on IN guinea NN pigs NNS suggest VVP that IN the DT TT NP dimer NN also RB triggers VVZ the DT tanning VVG response NN .
SPACE development

- expand the database (SPACE07 – SPACE12):
  - add new domains (psychology, medicine) and stratify according to publication time
to over 2 million words
- compare different cultures in terms of native and non-native, (sub-)discipline-specific variation
- apply in teaching at advanced university levels as a specialisation for language service providers
3. Concepts and illustrations

3.1 Complexity

higher complexity = more processing time → help the reader through commas, etc.
exemplary illustration in text excerpts, esp. headlines, lexical complexity
syntactic complexity: commas, words/sentence, etc.
Illustrating complexity: specialised vs. popular titles

0001AX  Indeterminate-length quantum coding
0001NS  The ultimate computer
0002AX  Quantum phase transitions and the breakdown of classical General Relativity
0002NS  What lies beneath
0046AX  The disruption of stellar clusters containing massive Black Holes near the galactic center
0046NS  Star shepherds
0104PN Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants
Phylogenetic relationships within Plantaginaceae were determined from a 4,730-nt data set consisting of portions of four chloroplast regions (ndhF, rbcL, and *intergenic spacers* *atpB rbcL* and *trnLtrnF*). Relationships within Plantago subgenus Plantago were analyzed from a 9,845-nt data set containing two additional chloroplast regions (*intergenic spacers* *psaAtrnS* and *trnCtrnD*). **Maximum likelihood (ML) trees** were constructed with PAUP* by using the general time-reversible model, a gamma distribution with four rate categories, and an estimate of the proportion of invariant sites. The rate matrix, base frequencies, shape of the **gamma distribution**, and proportion of invariant sites were estimated before the ML analysis from a neighbor joining tree constructed from the data. Divergence times outside Plantaginaceae were taken from ref. 27. Those within the family were calculated by using a penalized likelihood approach (28) as implemented in the R8S program (29) and a time constraint of 48 million years (27) for the Antirrhinum Plantago split. The ML tree was used as the starting tree for the divergence time analysis. The starting tree was constructed by first constraining the taxa in the 4,730-nt data set to incorporate the alternative relationships within subgenus Plantago and then estimating branch lengths for this topology in PAUP*. A smoothing factor of three was determined by using the R8S cross-validation procedure.

0104NS Plant DNA shows speedy changes
The **mitochondria** of a group of nondescript flowering plants contain the fastest-evolving DNA yet known. Until now, the **mitochondrial** genomes of plants were thought to evolve slowly. But when Jeffrey Palmer and colleagues at Indiana University in Bloomington compared mitochondrial DNA from nine species of plantain (members of the genus Plantago) and 41 other plants, they found that some Plantago sequences changed.
## Specialised vs. popular titles and commas

<table>
<thead>
<tr>
<th>text</th>
<th>title</th>
<th># commas</th>
<th>max. commas /sent.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001AX</td>
<td>Indeterminate-Length Quantum Coding</td>
<td>243</td>
<td>8</td>
</tr>
<tr>
<td>0001NS</td>
<td>The Ultimate Computer</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>0002AX</td>
<td>Quantum Transitions &amp; Breakdown of Relativity</td>
<td>110</td>
<td>8</td>
</tr>
<tr>
<td>0002NS</td>
<td>What Lies Beneath</td>
<td>113</td>
<td>4</td>
</tr>
<tr>
<td>0003AX</td>
<td>Mirror World Explanation for Pioneer spacecraft?</td>
<td>206</td>
<td>6</td>
</tr>
<tr>
<td>0003NS</td>
<td>Strange Attraction</td>
<td>129</td>
<td>6</td>
</tr>
<tr>
<td>0005AX</td>
<td>Realization of Quantum Games on Quantum Comp.</td>
<td>79</td>
<td>6</td>
</tr>
<tr>
<td>0005NS</td>
<td>Multiple Choice</td>
<td>134</td>
<td>5</td>
</tr>
<tr>
<td>0007AX</td>
<td>Hints of Gravity in Large Extra Dimensions?</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>0007NS</td>
<td>Pulling Power</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>0008AX</td>
<td>Astronomical Detection of Chromaticity Effects</td>
<td>301</td>
<td>9</td>
</tr>
<tr>
<td>0008NS</td>
<td>Spot the Stargate</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>0009AX</td>
<td>Cold Dark Matter from Dark Energy</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>0009NS</td>
<td>Darker and Darker</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>0010AX</td>
<td>Dimming Supernovae without Cosmic Acceleration</td>
<td>31</td>
<td>8</td>
</tr>
<tr>
<td>0010NS</td>
<td>Go-Faster Universe May Just Be a Trick</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>0011AX</td>
<td>Deuterium Burning in Jupiter Interior</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>0011NS</td>
<td>Twinkling Planet</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>0012AX</td>
<td>Computing a Glimpse of Randomness</td>
<td>447</td>
<td>12</td>
</tr>
<tr>
<td>0012NS</td>
<td>Smash and Grab</td>
<td>94</td>
<td>4</td>
</tr>
</tbody>
</table>
3.2 Hedges
complex, gradient, culture-specific

more than 60 'hedges and related phenomena', including
sort of, kind of, rather, basically, very, often, almost, as it were, in one sense, a regular, so to say, in name only, really, pseudo-, etc.

Brown/Levinson (1987: 145) "a particle, word or phrase that modifies the degree of membership of a predicate or a noun phrase in a set; it says of that membership that it is partial or true only in certain respects, or that it is more true and complete than perhaps might be expected".

Ventola/Mauranen (1996): Finns writing in English showed the tendency to stick to a few 'safe' expressions of epistemic modality, had less variation in the expressions than did native speakers of English, i.e. they did not behave in a native-like manner

→ def.: hedging = down-(up)-scaling author commitment (cf. stance, boosters)
3.2.1 Overlapping functional concepts

key terms: stance > hedging > modality

- “personal feelings, attitudes, value judgments, or assessments” (Biber et al. 1999: 966)
- “subjective” evaluation on the basis of own knowledge, experience, etc.
- context-dependency

academic culture in the discourse community determines how stance is expressed!
3.2.2 Formal indicators of hedging

standard example: modal auxiliaries
in epistemic use (Greek: “knowledge”)
in decreasing strength/propensity:
  *must*, *will*, *would*, *should*, *can*, *could*, *may*, *might*, ...

but also **modal adjuncts** / **adverbs** / **nouns**
<table>
<thead>
<tr>
<th>strong</th>
<th>assuredly</th>
<th>certainly 1</th>
<th>clearly 16</th>
<th>definitely</th>
<th>incontestably</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>indubitably</td>
<td>ineluctably</td>
<td>inescapably</td>
<td>manifestly</td>
<td>necessarily</td>
</tr>
<tr>
<td>obviously 4</td>
<td>patently</td>
<td>plainly</td>
<td>undoubtedly</td>
<td>surely 1</td>
<td>truly</td>
</tr>
<tr>
<td>unarguably</td>
<td>unavoidably</td>
<td>evidently</td>
<td>presumably</td>
<td>unquestionably</td>
<td>seemingly</td>
</tr>
<tr>
<td>ii</td>
<td>apparently 3</td>
<td>doubtless</td>
<td>evidently</td>
<td>presumably</td>
<td>seemingly</td>
</tr>
<tr>
<td>iii</td>
<td>arguably</td>
<td>likely 1</td>
<td>probably</td>
<td>possibly 1</td>
<td></td>
</tr>
<tr>
<td>iv</td>
<td>conceivably</td>
<td>maybe</td>
<td>perhaps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>weak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.3 Cohesion / coherence

3.3.1 Definitions

coherence = extralinguistic factors contributing to the creation of texture in the mind of the receiver!

cohesion = linguistic means which create texture
3.3.2 Formal indicators: cohesive devices

formal:
- conjunctions: *but, while*
- adverbs: *first, then, finally*

functional:
- sentence adverbials, e.g. clause-initial adverb  *ly,*
List 1: AntConc concordance for *ly, in the first few SPACE Corpus files

Subsequently, Braunstein et al. [5] presented a quantum analogue to classical Huffman coding. Because a general un
on in Section 2.3, below.) More recently, Chuang and Modi have developed a quantum version of arithmetic coding as a route to quantum data c
to be simply condensable codes. Obviously, all simply condensable codes are condensable, but the converse is not true. 4 The condensability
is that the code be prefix-free informally, that no initial segment of a code word is itself a code word. In the next section, we will show s
clusion of our quantum program.) Finally, the computer contains an output 
apel of qubits (initially all in the state |0\rangle on which the cond
of this register is called R_i,k. Initially, each register contains a set of code words from a fixed prefix-free quantum code. 13 Tape There is a ket
be made as small as possible. Conversely, in a simple condensation process, we must keep at least h_k qubits per signal to maintain high fid
elseless quantum coding theorem. Finally, we will show that the relative entropy is a measure of the additional resources (qubits) required t
onitions |1i of .1 and |2i of .2. (Equivalently, we can fix one of the purifications |1i and maximize over the other purification |2i.) The fidelity
aracteristic h_i < S_i) + 1. (59) Asymptotically, this code will achieve high fidelity using about S_i qubits per signal. An alternate scheme is e
ic bound to the code word length codewords themselves are stored immediately, before the remainder of the string is received [3]. But this terminology is inapplicable to the qua
with a new kind of quasiparticle. Astonishingly, these quasiparticles obey their own version of special relativity. For example, there's an absolute a
revolutions per minute (rpm), mena with fore and aft elements refer switch by ground command or r
from distances beyond 67 AU. Starting in the spring of 1996, s
sautaneous two-way range delay. Unfortunately, an instantaneous comparison was not possible in this case. The reason is that the signal-to-noise
m, 62, 60) at the Spain complex. Specifically, the Pioneers used (DSS 12, 14, 42, 40, 62, 60), Galileo used (DSS 12, 14, 42, 40, 63), and Ulysses C
resolution of the Resolver.1 Consequently, the JPL Doppler records are not frequency measurements. Rather, they are digitally counted measure
ance between these two points. Consequently, while studying the effect of a systematic error from propagation of the S-band carrier wave throu
onstant > 0.3 mm/s.) Consequently, the quoted errors are realistic, not formal, and represent our attempt to include systematics and a
orm on solar radiation pressure.) Finally, the parameter a(U) was determined by linear least squares. The best-fit value was obtained a(U)
the spin period decreases very quickly, while in between maneuvers the spin rate actually tends to increase at a rate of ~ (+0.0073±0.0003
of 28 maneuvers in all. As noted previously, in fitting the Pioneer 10 data over 11.5 years we used the standard space-fixed \( J_2 \) coordinate s
2s from their ULS counterparts. Nonetheless, acceleration offset, set
nally, there is the annual term. It remains in the data (for both Pioneers 10 and 11). A representation of
List 2: occurrences of unfortunately in SPACE07

1. It is unfortunately not the case that knowledge of the mechanical
2. Unfortunately, an instantaneous comparison was not possible
3. Unfortunately, exact information on gas usage is unavailable
4. Unfortunately, one has no a priori way of predicting
5. Unfortunately, neither easily works.
6. Unfortunately, there are as yet no reliable calculations that
7. Unfortunately, our experience is confined to an equilibrium ...
8. Unfortunately, the energy resolution of a neutrino telescope is ...
9. ... technique to obtain black hole masses which, unfortunately, is unfeasible
10. Unfortunately, lack of information about the collective ...
11. Unfortunately, the complexity of each subprocess also grows ...
12. Unfortunately, these theorems tell us practically nothing about ...
13. Unfortunately, there are very few analytic results available
14. Unfortunately, all recent experiments are, in principle,
15. Unfortunately, all those exciting recent experiments are
16. Unfortunately, those prior art solutions require daunting experiments
17. Unfortunately, the NID is uncomputable since the constituent ...
18. Unfortunately, in many places such information is getting harder ...
19. it's a good protective barrier," says Hildebrand, "unfortunately.
20. Unfortunately, because of the rarity of plant data from this ...

Definitions database concepts analysis teaching outlook 29/53
Undoubtedly the best-studied explanation, however, is...

... and so undoubtedly contributes something to Google pagecounts.

But one reaction would undoubtedly be near the top of both:

... undoubtedly protects stocks during periods of poor productivity.

... developing new cultivars ... would undoubtedly have an enormous impact.

Although these resources undoubtedly included plants, nearly all ...

... were eager to

List 3:occurrences of *undoubtedly* in SPACE07
List 3: occurrences of undoubtedly in SPACE07

1. Undoubtedly, the best-studied explanation, however, is ...
2. and so undoubtedly contributes something to Google pagecounts.
3. But one reaction would undoubtedly be near the top of both:
4. undoubtedly protects stocks during periods of poor productivity
5. developing new cultivars ... would undoubtedly have an enormous impact
6. Although these resources undoubtedly included plants, nearly all ...
7. Over time he acquired many imitators; undoubtedly some ... were eager to

modal and evaluative adverbs are not only used in popular academic English!
4. Analysis and quantification

- complexity: pop < spec
  - lexical
  - syntactic
- hedging: pop < spec
  - modal auxiliaries
  - frequency and propensity
- cohesion:
  - sentence adverbs
ComplexAna

Choose Input File

Browse...

Options
- Consider only one instance of a noun
- Use stoplist
- Reset coefficients to default values

Process Input File

Process I

Results

Number of tokens: 542
Number of words: 468
Maximum number of words in a sentence: 60
Mean number of words in a sentence: 19.500000
Number of nouns in text: 137
Number of nouns considered (not in stoplist): 87
Number of nouns considered & known to WordNet (%): 90.80%
Number of nouns considered & unknown to WordNet (%): 9.20%
Number of nouns considered & not in frequency list (%): 54.02%
Maximum length of a noun considered: 13
Mean length of a noun considered: 6.747126
Number of commas: 21
Maximum number of commas in a sentence: 3
Maximum Degree of Semantic Specialization of a noun: 12
Degree of Semantic Specialization of the text: 8.253165
Degree of Semantic Difficulty: 18.635404

Coefficients

0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00

Help me!  Exit
## 4.1 Calculating complexity: $\text{pop} < \text{spec}$

<table>
<thead>
<tr>
<th>complexity in PhysicsTexts</th>
<th>047 PN</th>
<th>047 NS</th>
<th>048 PN</th>
<th>048 NS</th>
<th>049 PN</th>
<th>049 NS</th>
<th>050 PN</th>
<th>050 NS</th>
<th>051 PN</th>
<th>051 NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of words / tokens</td>
<td>7177</td>
<td>2572</td>
<td>7118</td>
<td>403</td>
<td>5961</td>
<td>638</td>
<td>5493</td>
<td>172</td>
<td>3229</td>
<td>156</td>
</tr>
<tr>
<td>max. words in a sentence</td>
<td>87</td>
<td>50</td>
<td>87</td>
<td>44</td>
<td>166</td>
<td>38</td>
<td>84</td>
<td>26</td>
<td>97</td>
<td>31</td>
</tr>
<tr>
<td>mean words in a sentence</td>
<td>14.76</td>
<td>21.08</td>
<td>15.48</td>
<td>18.36</td>
<td>14.46</td>
<td>17.12</td>
<td>13.31</td>
<td>17.75</td>
<td>12.30</td>
<td>18.28</td>
</tr>
<tr>
<td>nouns in text</td>
<td>2346</td>
<td>725</td>
<td>2474</td>
<td>124</td>
<td>2076</td>
<td>149</td>
<td>1887</td>
<td>53</td>
<td>1145</td>
<td>53</td>
</tr>
<tr>
<td>nouns unknown to WordNet (%)</td>
<td>46.29</td>
<td>13.8</td>
<td>38.76</td>
<td>7.89</td>
<td>37.7</td>
<td>11.58</td>
<td>41.08</td>
<td>7.69</td>
<td>33.11</td>
<td>11.9</td>
</tr>
<tr>
<td>nouns not in frequency list (%)</td>
<td>76.98</td>
<td>52.19</td>
<td>75.84</td>
<td>56.58</td>
<td>75.41</td>
<td>53.68</td>
<td>76.63</td>
<td>56.41</td>
<td>76.86</td>
<td>66.67</td>
</tr>
<tr>
<td>max. length of noun</td>
<td>44</td>
<td>19</td>
<td>43</td>
<td>13</td>
<td>43</td>
<td>15</td>
<td>42</td>
<td>12</td>
<td>43</td>
<td>13</td>
</tr>
<tr>
<td>mean length of a noun</td>
<td>6.89</td>
<td>6.96</td>
<td>7.14</td>
<td>6.80</td>
<td>6.77</td>
<td>6.71</td>
<td>6.61</td>
<td>5.97</td>
<td>7.03</td>
<td>6.78</td>
</tr>
</tbody>
</table>
4.2 Hedges according to word-class in subcorpora

<table>
<thead>
<tr>
<th>Nouns</th>
<th>Verbs</th>
<th>Modal Verbs</th>
<th>Adverbs</th>
<th>Adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>970</td>
<td>1218</td>
<td>6654</td>
<td>1493</td>
<td>1962</td>
</tr>
<tr>
<td>798</td>
<td>1218</td>
<td>5592</td>
<td>5031</td>
<td>1962</td>
</tr>
<tr>
<td>220</td>
<td>561</td>
<td>3161</td>
<td>2953</td>
<td>1517</td>
</tr>
<tr>
<td>269</td>
<td>330</td>
<td>304</td>
<td>2860</td>
<td>321</td>
</tr>
<tr>
<td>11195</td>
<td>10782</td>
<td>11195</td>
<td>5592</td>
<td>1195</td>
</tr>
</tbody>
</table>

frequency per 100,000 words

quantum S  quantum P  astro S  astro P  particle S  particle P
Modal auxiliaries in biosciences subcorpora
Epistemic *can/could* and *may/might* in specific and popular sciences (per 100000 words)
## 4.3 Stance markers: POS and propensity

<table>
<thead>
<tr>
<th>Hedge form</th>
<th>POS Tag, Penn Treebank</th>
<th>Intuitive Degree Propensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>About</td>
<td>IN</td>
<td>6</td>
</tr>
<tr>
<td>Aim</td>
<td>NN</td>
<td>3</td>
</tr>
<tr>
<td>All</td>
<td>JJ, DT</td>
<td>10</td>
</tr>
<tr>
<td>Almost</td>
<td>RB</td>
<td>9</td>
</tr>
<tr>
<td>Approximating</td>
<td>VV</td>
<td>7</td>
</tr>
<tr>
<td>At least</td>
<td>JJS</td>
<td>9</td>
</tr>
<tr>
<td>Can</td>
<td>MD</td>
<td>5</td>
</tr>
<tr>
<td>Clearly</td>
<td>RB</td>
<td>9</td>
</tr>
<tr>
<td>Closest</td>
<td>JJS</td>
<td>7</td>
</tr>
<tr>
<td>Compared to</td>
<td>VV</td>
<td>5</td>
</tr>
<tr>
<td>Depending</td>
<td>VV</td>
<td>5</td>
</tr>
<tr>
<td>Dramatically</td>
<td>RB</td>
<td>8</td>
</tr>
<tr>
<td>Estimated</td>
<td>VV</td>
<td>7</td>
</tr>
<tr>
<td>Extremely</td>
<td>RB</td>
<td>9</td>
</tr>
</tbody>
</table>
## 4.4 Types of sentence adverbs in the SPACE07 Corpus

<table>
<thead>
<tr>
<th>readership adverb types</th>
<th>specialized</th>
<th>popular</th>
<th>relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AX</td>
<td>PN</td>
<td>AX+PN</td>
</tr>
<tr>
<td>evaluative</td>
<td>581</td>
<td>160</td>
<td>741</td>
</tr>
<tr>
<td>modal</td>
<td>36</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td>linking</td>
<td>34</td>
<td>129</td>
<td>163</td>
</tr>
<tr>
<td>domain-specific</td>
<td>131</td>
<td>42</td>
<td>173</td>
</tr>
</tbody>
</table>
5. Teaching: EAPspec - EAPpop writing courses

at M.A. level for language service providers
modules in English as a Global Language, TESOL
EAP spec as EAP theory & practice
EAP pop “science journalism” as part of journalistic writing

few models:
www.uefap.co.uk: low for “higher education”

→ get to know the community of practice by reading
5.1 Text analysis as awareness raising

popularisation features

- simplification
  - lexical
  - syntactic
- explicification
  - stance / hedging
  - cohesive devices
- personalisation
- sensationalism?
**Tumor cell surface heparan sulfate as cryptic promoters or inhibitors of tumor growth and metastasis**

Dongfang Liu*, Zachary Shriver*, Ganesh Venkataraman†, Yosuf El Shabrawi*, and Ram Sasisekharan*‡

*Division of Bioengineering and Environmental Health, †Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Robert Langer, Massachusetts Institute of Technology, Cambridge, MA, October 30, 2001 (received for review September 18, 2001)

**Heparan sulfate glycosaminoglycans**, present at the cell surface and in the extracellular matrix that surrounds cells, are important mediators of complex biological processes. Furthermore, it is now apparent that cells dynamically regulate the structure of their heparan sulfate “coat” to **differentially** regulate extracellular signals.

In the present study, the importance of sequence information contained within tumor cell-surface heparan sulfate was investigated. Herein, we demonstrate that the **heparan sulfate** glycosaminoglycan coat present on tumor cells contains bioactive sequences that impinge on tumor-cell growth and metastasis. **Importantly,** we find that growth promoting as well as growth inhibiting sequences are contained within the polysaccharide coat.

**Furthermore,** we find that the dynamic balance between these distinct polysaccharide populations regulates specific intracellular signal-transduction pathways.

[http://www.pnas.org/content/99/2/568.full](http://www.pnas.org/content/99/2/568.full)
Six blind men and the elephant—the many faces of heparan sulfate

Ajit Varki*

Author Affiliations

Glycobiology Research and Training Center, Departments of Medicine and Cellular and Molecular Medicine, University of California at San Diego, La Jolla, CA 92093-0687

It was six men of Indostan
To learning much inclined,
Who went to see the Elephant
(Though all of them were blind),
That each by observation
Might satisfy his mind… .
—John Godfrey Saxe, based on an Indian fable

Did you hear the one about the chemist, the biochemist, the molecular biologist, the physiologist, the physician-scientist, and the clinical oncologist? They went out to learn about a mysterious and complex creature called a heparan sulfate (HS) proteoglycan. Their interest was piqued by an intriguing paper in this issue of PNAS (1) with the title “Tumor Cell Surface Heparan Sulfate as Cryptic Promoters or Inhibitors of Tumor Growth and Metastasis.” They came upon a knowledgeable female scientist who told them that HS chains belonged to a class of long acidic sugar chains called glycosaminoglycans (GAGs), which usually are attached to cell surfaces via a core protein (2). The combination of one or more HS chains and a core protein is called an HS proteoglycan (Fig. 1).

http://www.pnas.org/content/99/2/543.long
AN ANTI-CANCER “smart cell” that uses nanotechnology to penetrate tumours and destroy them from the inside has been developed in America. The drug-packed “nanocell” has been successfully tested on two forms of cancer in mice, shrinking tumours and prolonging survival far more effectively than other therapies.

If the technique works as well in human beings, it could transform the treatment of many cancers, allowing doctors to harness nanotechnology to kill tumours without affecting healthy tissue.

At present, cancer is generally treated by surgery, or by chemotherapy and radiotherapy, which cause serious side-effects as they destroy healthy cells, too. The potential of nanotechnology to improve on this, creating drugs that attack cancer cells alone, excites many scientists.

Ram Sasisekharan, Professor of Biological Engineering at the Massachusetts Institute of Technology (MIT), who led the research, said: “The fundamental challenges in cancer chemotherapy are its toxicity to healthy cells and drug resistance by cancer cells.”

He said that nanocells promised to overcome both problems. In experiments on mice with melanoma, a form of skin cancer, and the Lewis variant of lung cancer, the nanocell proved considerably more effective than standard treatments.

Some 80 per cent of the mice with melanoma treated in this way survived for at least 65 days. This compared with survival of 30 days for mice treated with the best current therapy, and 20 days for untreated animals. There were also benefits against lung cancer, though these were less dramatic, which suggests that the technology will need to be adjusted to attack different forms of the disease.

http://www.timesonline.co.uk/tol/news/world/us_and_americas/article548769.ece
**NS: Pass the sugar**
26 January 2002 by Claire Ainsworth

THE sugar coating that surrounds our body's cells **could** be exploited to target tumours, by using fragments of the sugars to stop cancer cells in their tracks. Much of the research into cancer and other diseases has focused on DNA and proteins. But the long, straight sugar molecules on the surfaces of cells play a key role in how cells interact with their surroundings. "They are crucial to how cells filter information coming in from the outside," **says cell biologist Ram Sasisekharan.**

Normally cells maintain a delicate balance, changing their sugar coats only when their environment demands it. But **Sasisekharan thinks** tumour cells switch their coatings at whim, looking for combinations that enable them to grow and spread.

Sugars have been hard to study because they are highly complex. They have dozens of building blocks, compared with four for DNA, and unlike proteins, sugars don't have a fixed DNA blueprint.
Mail Online

The microscopic anti-cancer 'smart bomb' Last updated at 09:07 28 July 2005

Scientists have developed an anti-cancer "smart bomb" that can burrow into a tumour and detonate while leaving healthy cells unscathed. The drug-packed "nanocell" proved effective and safe against two distinct types of cancer in mice, it has been revealed. It mounts a two-pronged attack against cancer cells, by both cutting off their blood supply and destroying them with a toxic chemical agent. The approach can be compared with dropping a bomb on the enemy while at the same time cutting off its supply lines, say scientists at the Massachusetts Institute of Technology (MIT).

Tumour cells generate their own network of blood vessels to provide them with nutrients and oxygen through a process called angiogenesis. Many researchers are exploring the idea of preventing angiogenesis to starve tumours to death. But cutting off oxygen from cancer cells can prompt them to create new blood vessels and begin spreading.

An obvious solution is to combine anti-angiogenesis with chemotherapy, so that a tumour is destroyed before it has a chance to re-build its blood vessels.

However this kind of combination therapy faces an inherent problem. Cutting off the supply lines also removes the means by which chemotherapy drugs reach the tumour.

Professor Sasisekharan, who led the MIT research team in Cambridge, USA, said: "You can't deliver chemotherapy to tumours if you have destroyed the vessels that take it there.

http://www.dailymail.co.uk/health/article-357346/The-microscopic-anti-cancer-smart-bomb.html
5.2 Teaching guidelines

complexity:
- simplify sentence structures
- simplify complex NPs through hypernyms in an ontological hierarchy and modification

hedging:
- control all stance features carefully

coherence:
- from explicit to implicit cohesive features
5.3 Teaching exercises

- reading to get to know the discourse conventions: from surfing the corpus back to the originals including discussion boards, etc.

- find 2 types of titles:
  - simple and transparent NS titles vs. catchy boulevard titles

- re-writing / editing exercises
  - simplify lexical and syntactic structures
  - find transparent, explicit options

- writing popular academic texts and newspaper texts on the basis of specific academic texts
  - 1 sentence per paragraph
  - write freely first and reedit carefully later according to guidelines
5.4 MA module: “academic writing”

- course elements:
  - mini-corpus-compilation (portfolio + best practice)
  - structuring exercises (group work with discussion):
    - mark key phrases in argumentation
    - mark theme - rheme (Functional Sentence Perspective)
    - mind map of academic genres/text-types

- practice elements
  - placement/internship?
    - with language service provider or partner university?
  - assignments from other faculties?
6. **Outlook**

open up options for languages services
develop applied student skills in a wider context of conference / publishing projects
develop an MA module “academic writing” for a SE European university network (TEMPUS) with partners in Albania, Montenegro, Macedonia, etc.?
References


“Only connect ...”

- to your reader
- to your students
- to your partners
josef.schmied@phil.tu-chemnitz.de