Analysing and Teaching Englishes of Science

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
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.

http://en.wikipedia.org/wiki/Social_constructionism (23/6/10)

Constructivism applied to Englishes 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

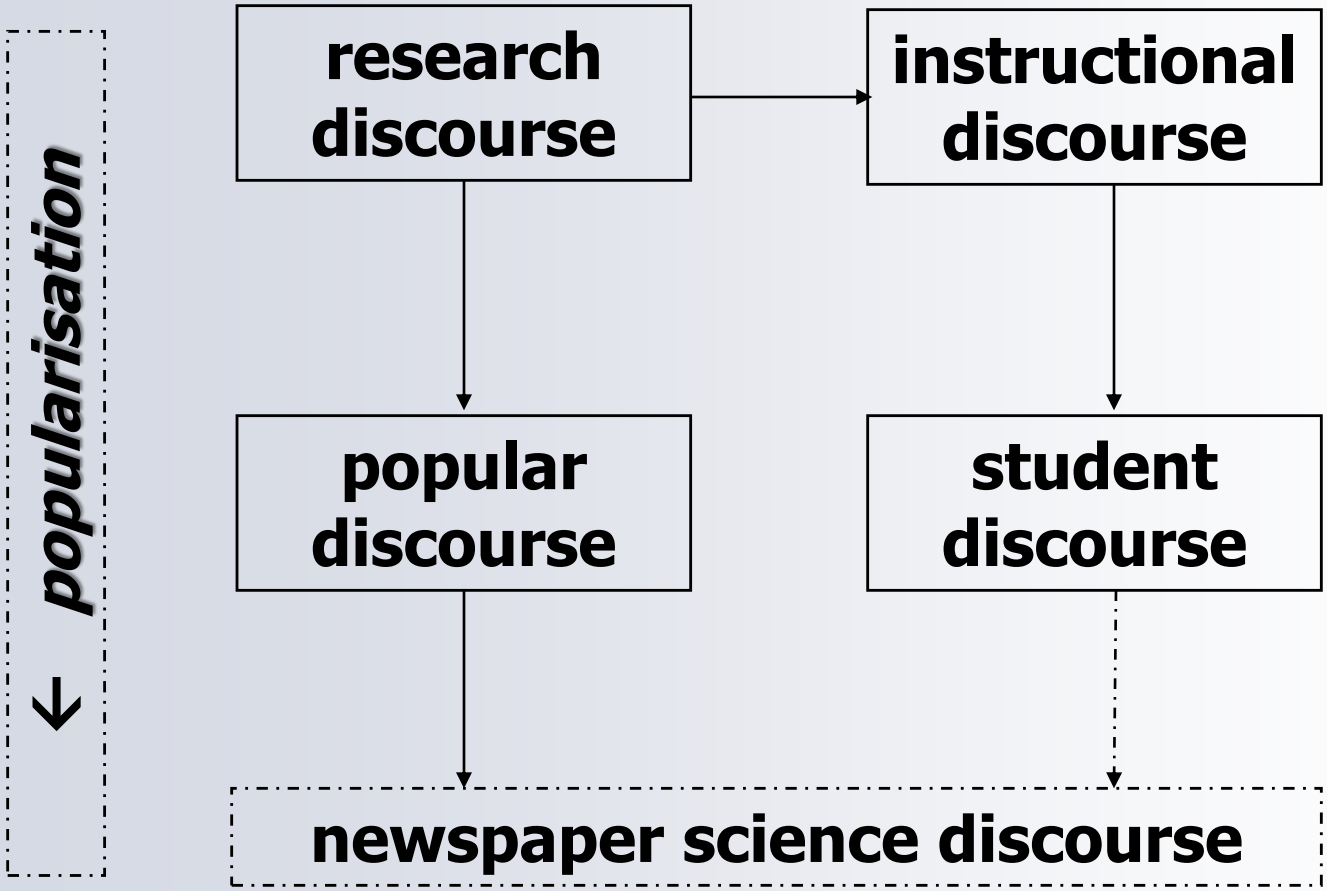
  "writing is always a personal and socio-cultural act of identity whereby writers both signal their membership in a range of communities as well as express their own creative presence" (Hyland 2006: 35)

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



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

student "literacy"
- **fieldwork notes/essays**
- **Mag/BA thesis**
- **seminar presentations**

science "journalism"
- **popular science articles**
- **popular science books**
- **newspaper science articles**

discipline-specific culture-specific

author-specific culture-specific

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

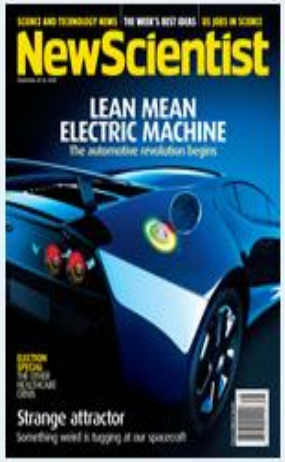
- science journals like *New Scientist* (subscription)
- academic online databases, pre-publication servers like *arXiv* (arxiv.org)
- publications in the *Proceedings of the National Academy of Sciences* (PNAS, pnas.org)
- plus: e.g. *Public Library of Science - Medicine* (plos.org)

NewScientist

SEARCH

PRINT EDITION

Subscribe



- Current issue
- Archive
- Full Access

NEW SCIENTIST JOBS

NEW SCIENTIST STUDY

SUBSCRIPTIONS CENTRE

Get 4 extra **free issues** and unlimited **free access** to NewScientist.com

SUBSCRIBE

RENEW

GIFT SUBSCRIPTION

TOP STORY

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?

BREAKING NEWS

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

VIDEO: BEST OF THE WEEK

more video

BLOG: GENETICS

Google genes

Sergey Brin has a mutation linked to Parkinson's

PHOTO GALLERY

Coral quest

An expedition to islands off the Great Barrier Reef turns up hundreds of new species

PASSWORD LOGIN

username :

password:

Your login is case-sensitive

Log in

- Forgotten your password?
- Subscriber? Register now!
- Subscribe now
- Institutional Subscribers
- Athens login

PNAS

Proceedings of the National Academy of Sciences of the United States of America

Archive of All Online Issues

January 1915 – Present

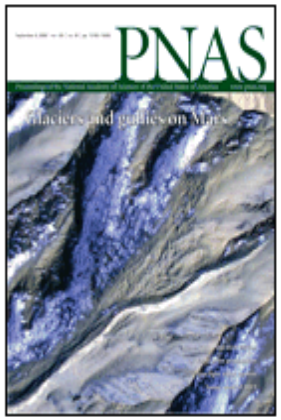
Collected Papers: List of Inaugural Articles || List of Commentaries || List of Reviews || List of Perspectives || List of Colloquia Papers || From the Academy

See also: PNAS Supplements Online and The Cover Archive

Current Issue:

September 16, 2008
Vol. 105, Num. 37

Recent Issues:



September 9, 2008
Vol. 105, Num. 36

September 2, 2008
Vol. 105, Num. 35

August 26, 2008
Vol. 105, Num. 34

Full Text and Abstracts: November 1996 – Present

2000s | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | –

About Our New Site Design >>

Search PNAS GO

advanced search >>

This Week's Issue

September 16, 2008, 105 (37)

From the Cover

- Evolution of the Web
- Neanderthal brain growth
- Mystery of crop origins
- Recipes for pilus assembly
- Bisphenol A disrupts synapses

Alert me to new issues of PNAS

Early Edition

arXiv.org

Open access to 497,864 e-prints in Physics, Mathematics, Computer Science, Quantitative Biology and Statistics

Subject search and browse: Physics Search Form Interface Catchup

Computer Science
Mathematics
Nonlinear Sciences
Physics
Quantitative Biology
Statistics

26 Jun 2008: Increased size
12 Mar 2008: NIH public acc
See cumulative "What's New
Robots Beware: indiscrimin ds from this site are *not* permitted.

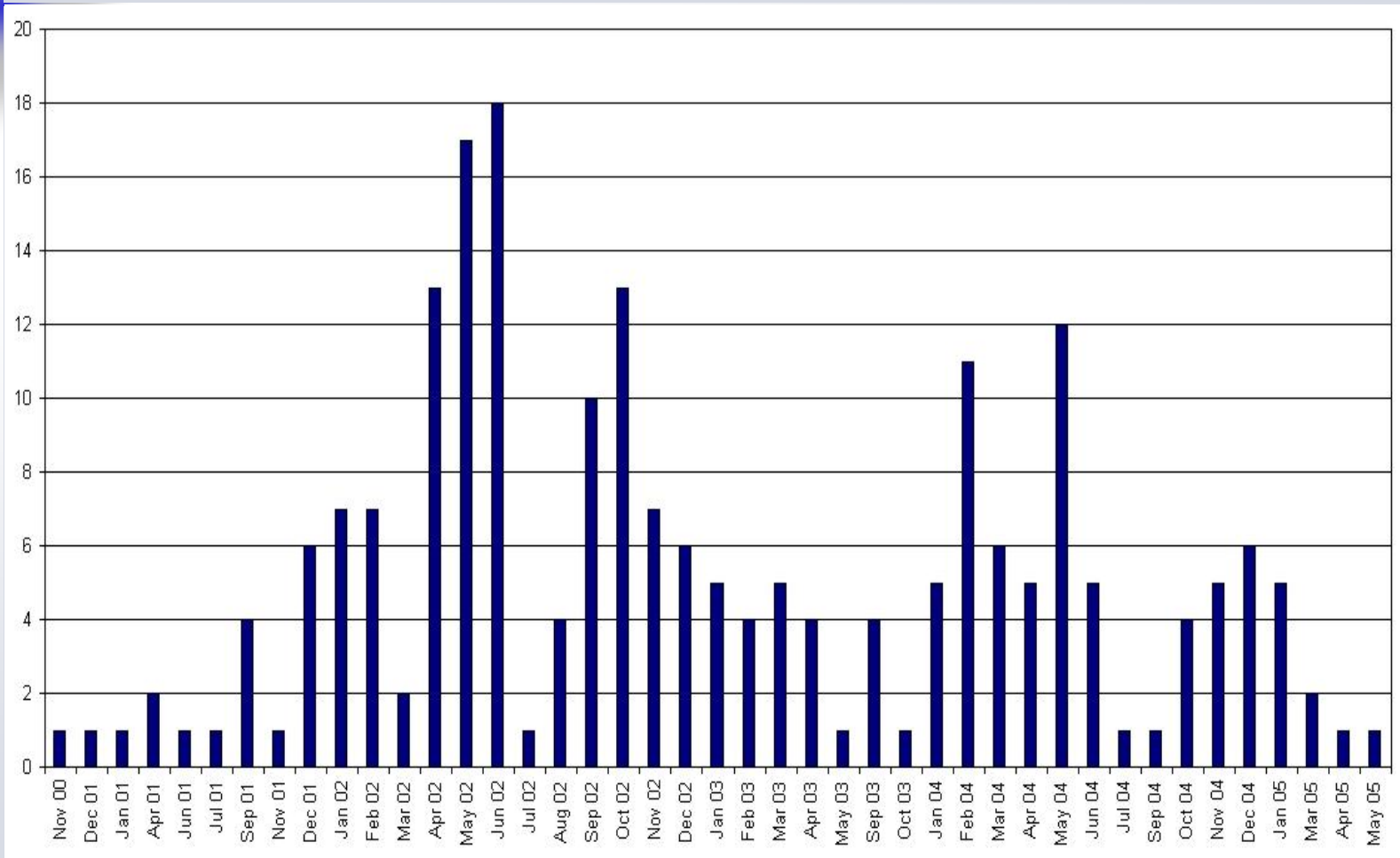
Physics

- Astrophysics (**astro-ph** new, recent, find)
- Condensed Matter (**cond-mat** new, recent, find)
  includes: Disordered Systems and Neural Networks; Materials Science; Mesoscopic Systems and Quantum Hall Effect; Other; Soft Condensed Matter; Statistical Mechanics; Strongly Correla
  Superconductivity
- General Relativity and Quantum Cosmology (**gr-qc** new, recent, find)
- High Energy Physics - Experiment (**hep-ex** new, recent, find)
- High Energy Physics - Lattice (**hep-lat** new, recent, find)
- High Energy Physics - Phenomenology (**hep-ph** new, recent, find)
- High Energy Physics - Theory (**hep-th** new, recent, find)
- Mathematical Physics (**math-ph** new, recent, find)
- Nuclear Experiment (**nucl-ex** new, recent, find)
- Nuclear Theory (**nucl-th** new, recent, find)
- Physics (**physics** new, recent, find)
  includes (see detailed description): Accelerator Physics; Atmospheric and Oceanic Physics; Atomic Physics; Atomic and Molecular Clusters; Biological Physics; Chemical Physics; Classical P
  Computational Physics; Data Analysis, Statistics and Probability; Fluid Dynamics; General Physics; Geophysics; History of Physics; Instrumentation and Detectors; Medical Physics; Optics; P
  Education; Physics and Society; Plasma Physics; Popular Physics; Space Physics
- Quantum Physics (**quant-ph** new, recent, find)

Table 1: Domain set-up of the SPACE10 corpus

readership / texts		specialised	popular	popular: special
		words	words	percentage
physics		**AX**	**NS**	
quantum-	11	115981	15327	13%
particle-	2	26384	4574	17%
astro-	9	167553	18119	11%
bioscience		**PN**	**NS**	
biochemistry	6	53212	7114	13%
genetics	24	183312	19101	10%
microbiology	10	53139	5176	10%
∑	62	599581	69411	12%
average		99930	11568	

Table 2: SPACE07 text publication dates



PN and NS SPACE texts tagged with POS and propensity value for hedges (0 no – 10 all)

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 data base (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 of language service providers

3. Concepts and illustrations

3.1 Complexity

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 sheperds*

Specialised vs. popular texts: lexical choices

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 MLanalysis 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

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

3.2 Hedges
complex, gradient, culture-specific

G. Lakoff (1972) "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts":

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)

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 3: A sample of modal adjuncts (from Huddleston/Pullum 2002:768)

strong					
i	assuredly	certainly 1	clearly 16	definitely	incontestably
	indubitably	ineluctably	inescapably	manifestly	necessarily
	obviously 4	patently	plainly	surely 1	truly
	unarguably	unavoidably	undeniably	undoubtedly 7	unquestionably
ii	apparently 3	doubtless	evidently	presumably	seemingly
iii	arguably	likely 1	probably		
iv	conceivably	maybe	perhaps	possibly 1	
weak					

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 **/y,*

Schmied
TALC9
Brno
030710

List 1: AntConc concordance for *ly, in the first few SPACE Corpus files

KWIC	File
of this idea several years ago. Subsequently, Braunstein et al. [5] presented a quantum analogue to classical Huffman coding. Because a general u	0001A
on in Section 2.5, below.) More recently, Chuang and Modha have developed a quantum version of arithmetic coding as a route to quantum data c	0001A
to be simply condensable codes. Obviously, all simply condensable codes are condensable; but the converse is not true. 4 The condensability	0001A
is that the code be prefix-free informally, that no initial segment of a zef codeword is itself a codeword. In the next section, we will show	0001A
cution of our quantum program.) Finally, the computer contains an output tape of qubits (initially all in the state |0i) on which the cond	0001A
f this register is called Ri,k. Initially, each register contains a zef codeword from a fixed prefixfree quantum code. 13 Tape There is a t	0001A
can be made as small as desired. Conversely, in a simple condensation process, we must keep at least hli qubits per signal to maintain high fid	0001A
iseless quantum coding theorem. Finally, we will show that the relative entropy is a measure of the additional resources (qubits) required t	0001A
tions |li of .1 and |2i of .2. (Equivalently, we can fix one of the purifications |li and maximize over the other purification |2i.) The fidelity	0001A
satisfies hli < S(.) + 1. (59) Asymptotically, this code will achieve high fidelity using about S(.)+1 qubits per signal. An alternate scheme is	0001A
ic bound to the codeword length exactly, without resorting to block coding? In other words, for what codes and codeword ensembles can we hav	0001A
codewords themselves are stored separately, in entangled strings of qubits. This means that the average number of qubits used to store the qua	0001A
vidual codeword from the string immediately, before the remainder of the string is received [3]. But this terminology is inapplicable to the qua	0001A
ith a new kind of quasiparticle. Astonishingly, these quasiparticles obey their own version of special relativity. For example, there's an absolute	0002N
8 revolutions per minute (rpm), respectively, with the spin axes running through the centers of the dish antennae. Their spin-stabilizations and	0003A
enna with fore and aft elements respectively, provided broad-angle communications at intermediate and short ranges. For DSN acquisition, these t	0003A
sfer switch by ground command or automatically, if needed. There is a redundancy in the communication systems, with two receivers and two transmi	0003A
le from distances beyond 67 AU. Recently, support of the Pioneer spacecraft has been on a non-interference basis to other NASA projects. It w	0003A
tarting in the spring of 1996. Currently, two types of Galileo navigation data are available, namely Doppler and range measurements. As ment	0003A
tantaneous twoway range delay. Unfortunately, an instantaneous comparison was not possible in this case. The reason is that the signal-to-noise	0003A
, 62, 63) at the Spain complex. Specifically, the Pioneers used (DSS 12, 14, 42, 43, 62, 63), Galileo used (DSS 12, 14, 42, 43, 63), and Ulysses	0003A
resolution of the resolver. Consequently, the JPL Doppler records are not frequency measurements. Rather, they are digitally counted measure	0003A
tance between these two points. Correspondingly, ril , ri2 , and ril2 are similar distances relative to a particular i-th body in the solar system	0003A
term are hard to quantify. But luckily, its effect on Doppler observables and, therefore, on our results is small. (We will address this is	0003A
nd in the literature [59]-[62]. Consequently, while studying the effect of a systematic error from propagation of the S-band carrier wave throug	0003A
nd Chandler wobble are obtained observationally, by means of Lunar and Satellite Laser Ranging (LLR and SLR) techniques and VLBI. Previously they w	0003A
different constant batch sizes; namely, 0, 5, 30, and 200 day batch sizes. (Later we also used 1 and 10 day batch sizes.) In each batch on	0003A
ers antennae [73]). As stated previously, the analyses were modeled to include the effects of planetary perturbations, radiation pressure, t	0003A
teristic .2 value of 0.3 mm/s.] Consequently, the quoted errors are realistic, not formal, and represent our attempt to include systematics and a	0003A
on on solar radiation pressure.) Finally, the parameter aP(U) was determined by linear least squares. The best fit value was obtained aP(U)	0003A
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	0003A
f 28 maneuvers in all. As noted previously, in fitting the Pioneer 10 data over 11.5 years we used the standard space-fixed J2000 coordinate s	0003A
2s from their WLS counterparts. Finally, there is the annual term. It remains in the data (for both Pioneers 10 and 11). A representation of	0003A
nomalous acceleration off- set. Mathematically, this is saying that in any interval i = I, II, III, for which the spin-rate change is an approximat	0003A

List 2: occurrences of *unfortunately* in SPACE07

1	It is unfortunately not the case that knowledge of the mechanical	0002AX
2	Unfortunately, an instantaneous comparison was not possible	0003AX1
3	Unfortunately, exact information on gas usage is unavailable	0003AX1
4	Unfortunately, one has no a priori way of predicting	0003AX1
5	Unfortunately, neither easily works.	0003AX1
6	Unfortunately, there are as yet no reliable calculations that	0004AX1
7	Unfortunately our experience is confined to an equilibrium …	0004AX2
8	Unfortunately, the energy resolution of a neutrino telescope is …	0015AX
9	… technique to obtain black hole masses which, unfortunately, is unfeasible	0018AX
10	Unfortunately, lack of information about the collective …	0021AX
11	Unfortunately, the complexity of each subprocess also grows …	0025AX
12	Unfortunately these theorems tell us practically nothing about …	0032AX
13	Unfortunately there are very few analytic results available	0040AX
14	Unfortunately, all recent experiments are, in principle,	0043AX
15	Unfortunately, all those exciting recent experiments are	0043AX
16	Unfortunately, those prior art solutions require daunting experiments	0043AX
17	Unfortunately, the NID is uncomputable since the constituent …	0045AX
18	Unfortunately, in many places such information is getting harder …	0056NS
19	it's a good protective barrier," says Hildebrand, "unfortunately."	0069NS
20	Unfortunately, because of the rarity of plant data from this …	0100PN

List 3: occurrences of *undoubtedly* in SPACE07

1	Undoubtedly the best-studied explanation, however, is …	0016AX
2	and so undoubtedly contributes something to Google pagecounts.	0045AX
3	But one reaction would undoubtedly be near the top of both:	0047NS
4	, undoubtedly protects stocks during periods of poor productivity	0059PN
5	developing new cultivars … would undoubtedly have an enormous impact	0087PN
6	Although these resources undoubtedly included plants, nearly all …	0100PN
7	Over time he acquired many imitators; undoubtedly some … were eager to	0102PN

List 3: occurrences of *undoubtedly* in SPACE07

1	Undoubtedly the best-studied explanation, however, is …	0016AX
2	and so undoubtedly contributes something to Google pagecounts.	0045AX
3	But one reaction would undoubtedly be near the top of both:	0047NS
4	, undoubtedly protects stocks during periods of poor productivity	0059PN
5	developing new cultivars … would undoubtedly have an enormous impact	0087PN
6	Although these resources undoubtedly included plants, nearly all …	0100PN
7	Over time he acquired many imitators; undoubtedly some … were eager to	0102PN

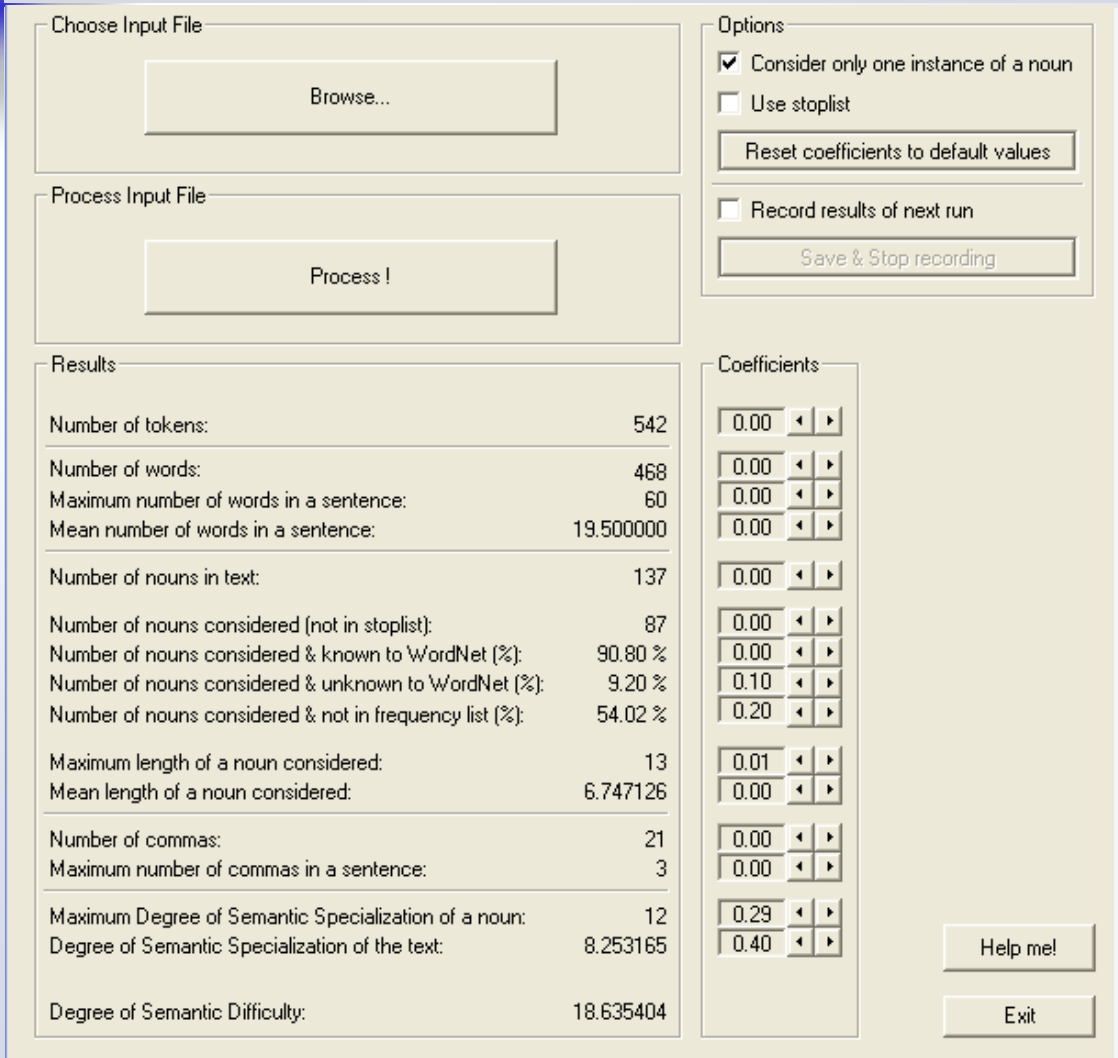
modal and evaluative adverbs are not only used in popular academic English!

4. Analysis and quantification

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

Schmied
TALC9
Brno
030710

ComplexAna

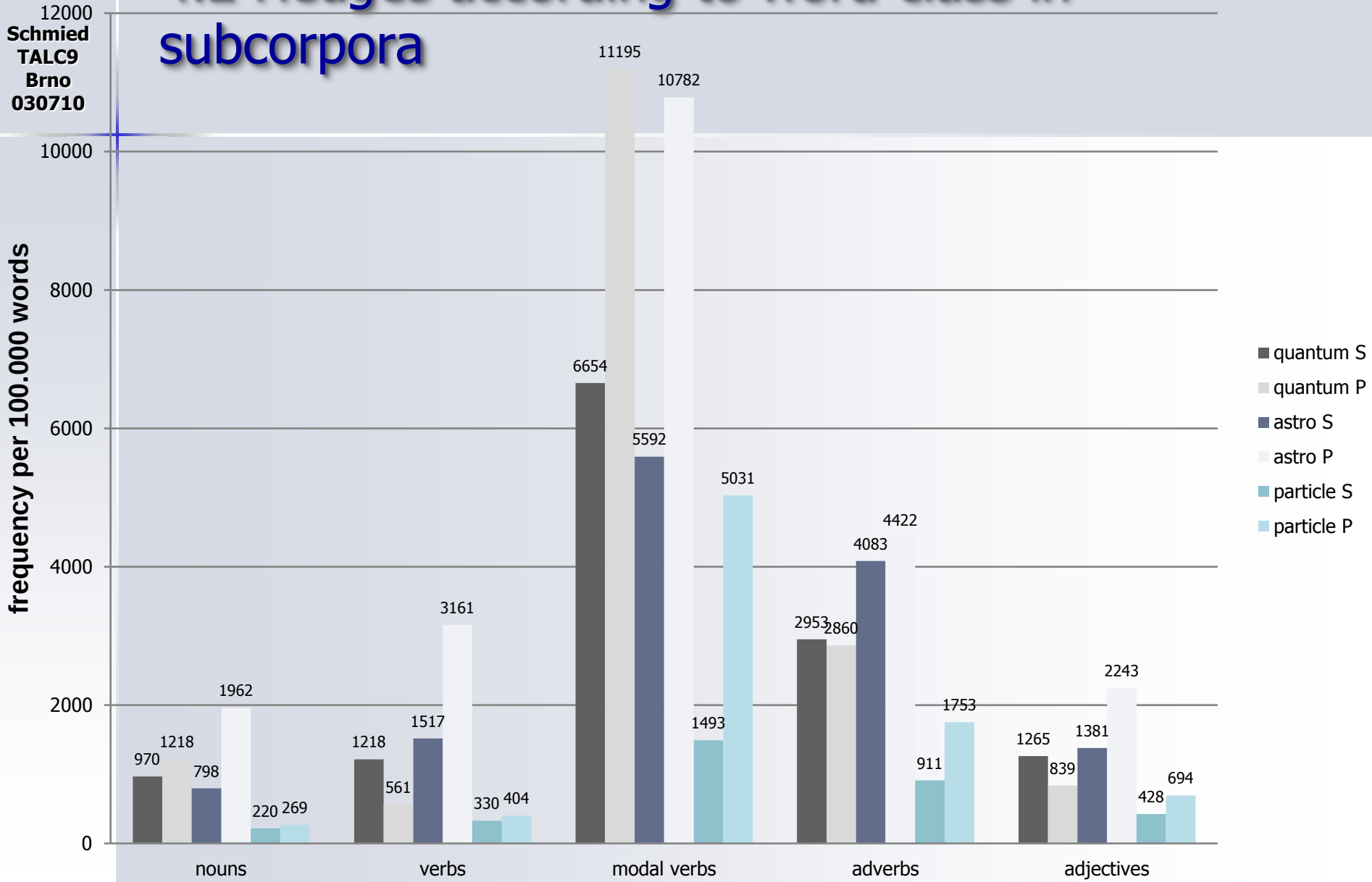


4.1 Calculating complexity : pop < spec

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

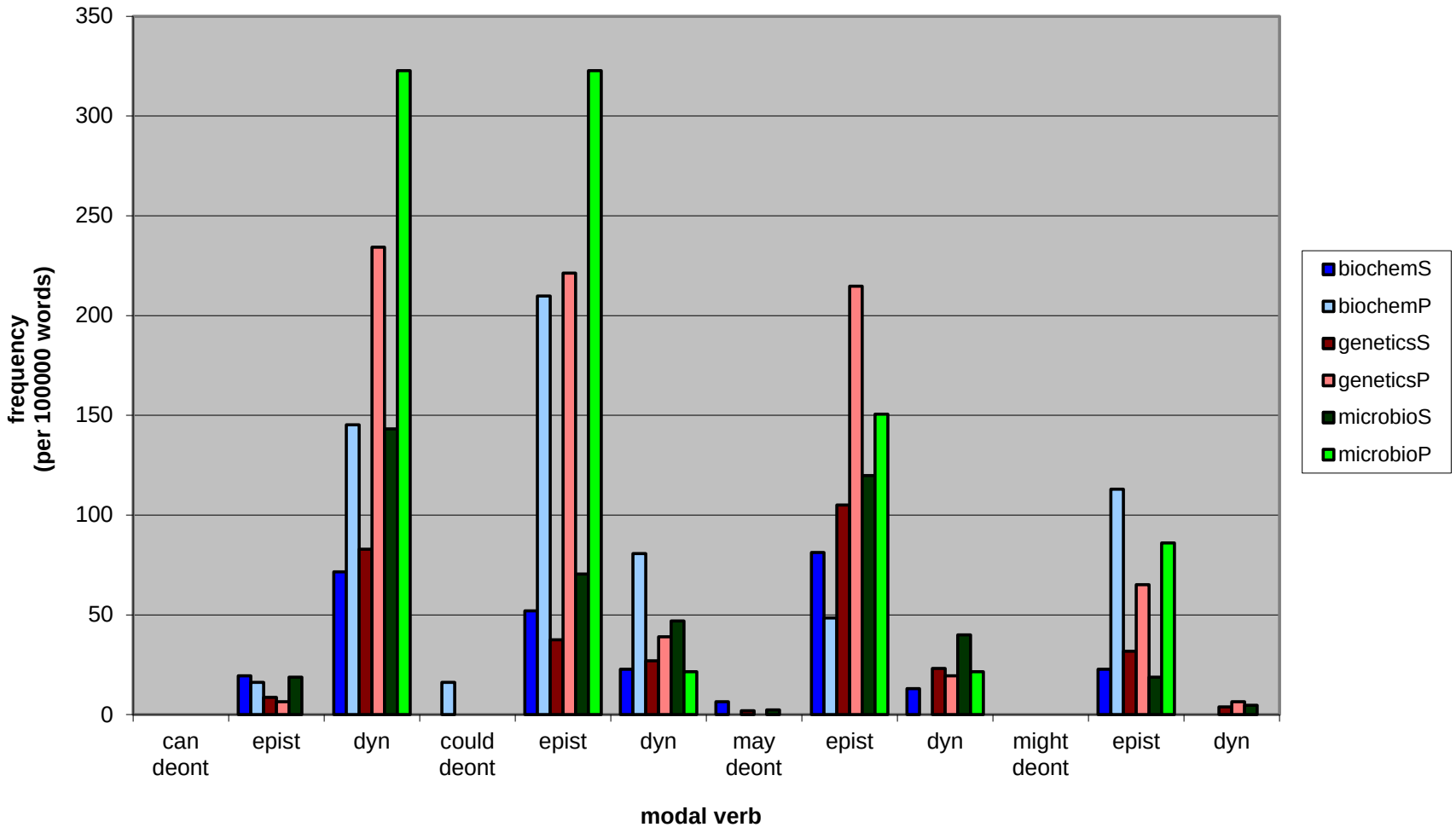
4.2 Hedges according to word-class in subcorpora

Schmied
TALC9
Brno
030710

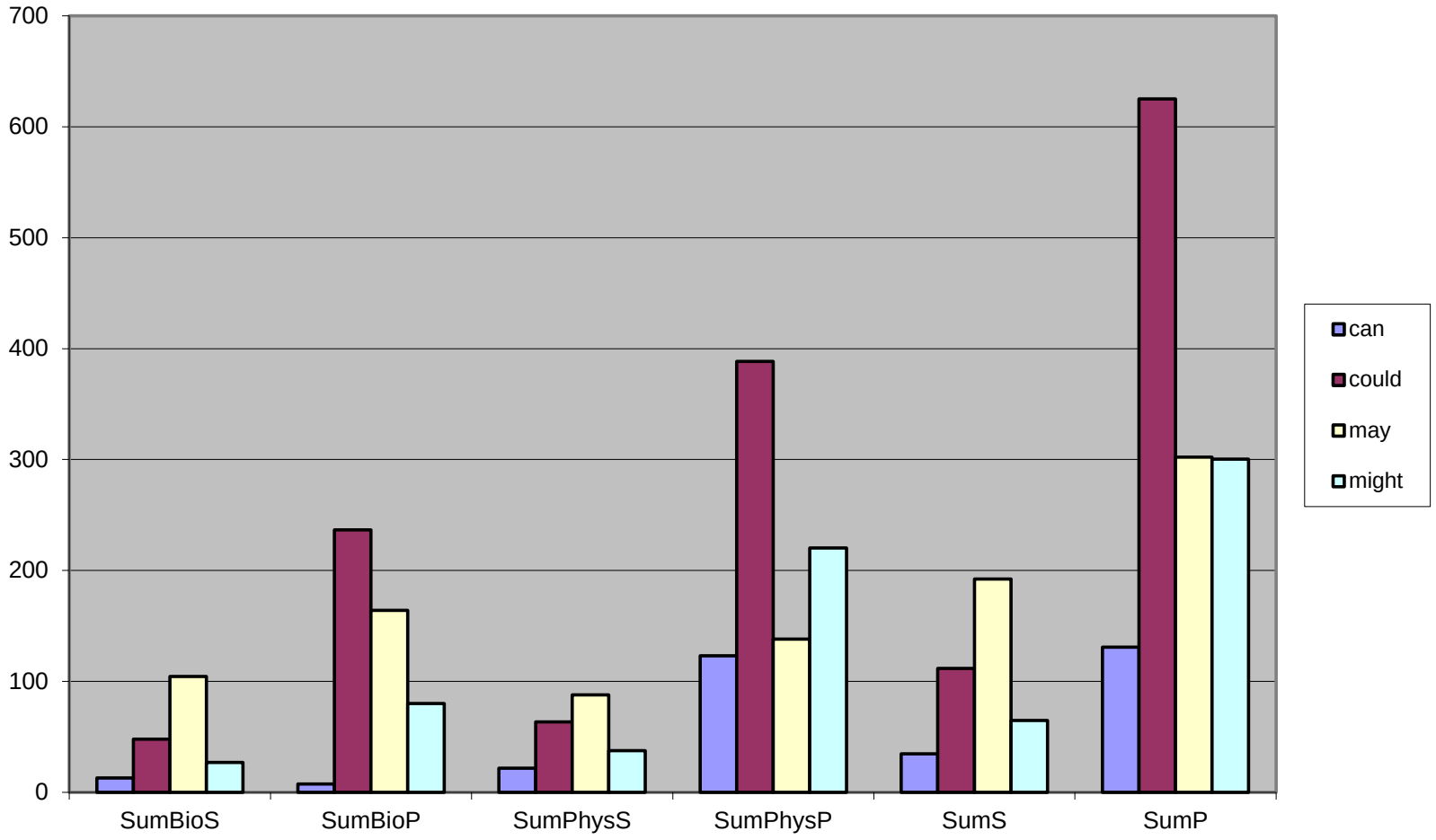


Schmied
TALC9
Brno
030710

Modal auxiliaries in biosciences subcopora



Epistemic *can/could* and *may/might* in specific and popular sciences (per 100000words)



4.3 Stance markers: POS and propensity

Hedge form	POS Tag, Penn Treebank	Intuitive Degree Propensity
About	IN	6
Aim	NN	3
All	JJ, DT	10
Almost	RB	9
Approximating	VV	7
At least	JJS	9
Can	MD	5
Clearly	RB	9
Closest	JJS	7
Compared to	VV	5
Depending	VV	5
Dramatically	RB	8
Estimated	VV	7
Extremely	RB	9

4.4 Types of sentence adverbs in the SPACE07 Corpus

readership adverb types	specialized			popular	relationship
	AX	PN	AX+PN	NS	NS:(AX+PN)
evaluative	581	160	741	69	9.30
modal	36	14	50	11	22
linking	34	129	163	8	4.90
domain-specific	131	42	173	3	1.70

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

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

From the Times
July 28, 2005

Cancer drug 'smart cell' can attack tumours from inside

By Mark Henderson, Science Correspondent

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

6. Outlook

develop applied student skills in a wider context of
conference / publishing projects

open up options for languages services

add a new dimension to TaLC:
EFL, ELF, ESL, EAP/ESP

References

Biber et al. (1999). *Longman grammar of spoken and written English.* London: Longman.

Brown, G. & S.C. Levinson (1987). *Politeness. Some universals in language usage.* Cambridge: Cambridge University Press.

Hyland , K. (2005). "Stance and Engagement: A Model of Interaction in Academic Discourse." *Discourse Studies: An Interdisciplinary Journal for the Study of Text and Talk* 2005 Apr 7 (2): 173-92.

Hyland, K. (2006). "The 'Other' English: Thought on EAP and academic writing". *The European Messenger 15(2), 34-39.*

Hyland, Ken (2009). *Academic Discourse: English in a Global Context*. London: Continuum.

Lakoff, G. (1972). "Hedges: A study in meaning criteria and the logic of fuzzy concepts." In Perenteau, P.M., Levi, J.N., Phares, G.C. *Papers from the Eighth Regional Meeting,* Chicago: Chicago Linguistic Society.

Swales (1990). *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.

Ventola, E. & A. Mauranen (eds. 1996). *Academic writing. Intercultural and textual issues.* Amsterdam: John Benjamins.