English as an International Academic Language: Exploring conventions in different disciplines

English world-wide/Africa + corpus/computer applications + academic writing

Josef Schmied
English Language & Linguistics
Chemnitz University of Technology
http://www.tu-chemnitz.de/phil/english/ling/presentations_js.php
josef.schmied@phil.tu-chemnitz.de
1. Introduction: Knowledge & cultures

1.1. Defining culture

James Spradley:
Culture is the acquired knowledge people use to interpret experience and generate behavior.

Where does “cultural variation” (in metadiscourse) come from in early academic knowledge presentation?

Contextualisation by author as
male - female?
language: first - second?
institutional acculturation in the discipline? on a functional basis?
1.2. Culture “with a small c”

People may use the word culture to mean any of these:

- **National / ethnic culture:**
  The group assumed to be site of child's primary socialization -- "THE Latvian culture," "THE African-American culture." In the U.S. this is usually the assumed meaning of culture, and people revert to this narrow view of culture out of habit, even when you may have been quite explicit about defining culture more broadly than this.

- **Secondary or subgroup culture:**
  cultural groups we've been socialized into: Organizational culture, professional culture, manager culture, Muslim culture, peer culture, prison culture, nerd culture and so on.

- **Culture in the anthropological sense:**
  the meanings and behaviors groups of people develop and share over time.

- **Capital C Culture:**
  the high arts of theater, painting, music, etc., or a superior upbringing.

http://www.culture-at-work.com/concept1.html (03/10/12)
1.3. Cultures in training academic knowledge processing

In academic knowledge presentation, culture may be included in the following variables:

- **gender**: male vs. female culture
  - a long research history in sociolinguistics

- **language**: English vs. Afrikaans vs. African Languages (Zulu/Sotho) culture
  - more diverse in English than other languages; *lingua franca*?

- **discipline**: science vs. humanities culture, etc.
  - long research tradition in EAP/genre analysis

trained conventions on a functional basis?
1.4. Culture in a genre approach to academic discourses

**Research “output”**
- **Research article**
  - Book reviews
  - Project proposals
- Conference presentations

**Science “journalism”**
- **Popular science articles**
  - Popular blogs (David Crystal)
  - Popular science films (Horizon)
- Popular science books

**Teacher “talk”/ e-learning**
- Ppt presentations
  - Lectures
  - Student presentations
- Textbooks
  - Wikis
  - www pages

**Student “literacy”**
- Fieldwork notes / essays
- MA/ BA thesis
- Seminar presentations

“Novice Academic English”
- Author-specific
- Culture-specific
2. The database:
ZAMA Corpus from Stellenbosch

- 66 texts
- gender balanced
- ethnic/language background according to name:
  English, Afrikaans, (S)African languages, rest (Indian, Chinese, Luo)
- department-stratified: aim 5 (6) MA theses in each of 13 disciplines
  - Anthropology
  - Chemistry
  - English (Lit.)
  - Curriculum Studies
  - Geography
  - History
  - Journalism
  - (Gen.) Linguistics
  - Mechanical Engineering
  - Civil Engineering
  - Psychology
  - (Public) Law
  - Political Science
SUNScholar is an open access electronic archive for the collection, preservation and distribution of digital materials created by members of Stellenbosch University, which are of great academic research importance. SUNScholar is built using open standards with open source software for the purposes of extremely long term digital preservation. SUNScholar is an online academic research information service of the library such as SUNJournals and SUNConferences.

Please read the help guides on our wiki before using this repository.

Wählen Sie einen Bereich, um die enthaltenen Sammlungen zu durchstöbern.

- Faculty of AgriSciences
- Faculty of Arts and Social Sciences
- Faculty of Economic and Management Sciences
- Faculty of Education
- Faculty of Engineering
- Faculty of Law
- Faculty of Medicine and Health Sciences
- Faculty of Military Sciences
- Faculty of Science
- Faculty of Theology
- General Collections
3. Text variation based on ComplexAna
3.1. **Text variation: tokens by gender**

![Bar chart showing mean number of tokens by gender.](chart.png)

- **Female**: Approximately 60,000 tokens
- **Male**: Approximately 40,000 tokens
- **Total**: Approximately 50,000 tokens
3.1. **Text variation: tokens by language**

![Bar graph showing number of tokens by language](image-url)
3.1. **Text variation: tokens by discipline**

**Mean number of tokens**

![Graph showing mean number of tokens by discipline](Image)
3.1. Text variation: tokens by year

Mean number of tokens

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean number of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>
3.2. Text complexity: WordNet nouns by gender

Mean of number of nouns known to WordNet (%)

- Female: 64%
- Male: 59%
- Total: 57%
3.2. Text complexity: WordNet nouns by language

Mean number of nouns known to WordNet (%)
3.2. Text complexity: WordNet nouns by discipline

Mean number of nouns known to WordNet (%)
3.2. Text complexity: Wordnet nouns by year

Mean of number of nouns known to WordNet (%)
4. Variation in personal pronoun usage

author involvement: I, me, my

reader address: you, your

semantic differences:
inclusive - exclusive we, pluralis majestatis/modestiae
you as audience - (any)one

examples:

„I will now review some of literature relevant to South Africa in order to provide a context for the study.” (ZMA06BA-YE)

„[…] these larger social patterns […] will have to be addressed if we are to curb this phenomenon.” (ZMA08MV-YX)

„Furthermore, you need to ‘comprehend ‘ [what] you’re reading […]” (ZMA10MM-EE)
### 4.1. Variation in personal pronoun usage

<table>
<thead>
<tr>
<th></th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; sing</td>
<td>6105</td>
<td>2112</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; pl</td>
<td>1427</td>
<td>844</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; sing</td>
<td>3200</td>
<td>419</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Afrikaans</th>
<th>English</th>
<th>S Afr Lang</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; sing</td>
<td>2540</td>
<td>6934</td>
<td>2867</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; pl</td>
<td>918</td>
<td>1313</td>
<td>1480</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; sing</td>
<td>925</td>
<td>3930</td>
<td>726</td>
</tr>
</tbody>
</table>
4.1. Variation in personal pronoun usage by gender

relative/1M words and by 4 functional categories
4.1. Variation in personal pronoun usage by language

relative/1M words and by 4 functional categories
### 4.2. Variation in personal pronoun usage by discipline

<table>
<thead>
<tr>
<th></th>
<th>Anth</th>
<th>Chem</th>
<th>EngLit</th>
<th>Geo</th>
<th>Hist</th>
<th>Jour</th>
<th>Ling</th>
<th>MechE</th>
<th>CivE</th>
<th>Pol</th>
<th>Curr</th>
<th>Law</th>
<th>Psych</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st s.</td>
<td>4272</td>
<td>904</td>
<td>9695</td>
<td>1132</td>
<td>1872</td>
<td>1347</td>
<td>5989</td>
<td>815</td>
<td>598</td>
<td>917</td>
<td>5998</td>
<td>386</td>
<td>11212</td>
</tr>
<tr>
<td>1st pl</td>
<td>1355</td>
<td>739</td>
<td>1710</td>
<td>264</td>
<td>1881</td>
<td>1179</td>
<td>1337</td>
<td>19</td>
<td>36</td>
<td>932</td>
<td>1872</td>
<td>73</td>
<td>2398</td>
</tr>
<tr>
<td>2nd pl</td>
<td>509</td>
<td>67</td>
<td>6163</td>
<td>481</td>
<td>421</td>
<td>839</td>
<td>3409</td>
<td>51</td>
<td>63</td>
<td>104</td>
<td>1319</td>
<td>0</td>
<td>4661</td>
</tr>
</tbody>
</table>
4.2. Variation in personal pronoun usage by discipline
5. Variation in cohesive devices: linkers

based on 50 linguistic variables (conjuncts, adverbs) in 4 functional categories

eamples:
„Also, within the context of years of negligence by Dutch officials, they deemed it their right to oversee their own destiny.” (ZMA06RHM-HE)

„Private contractors, rather than the military, benefited from the outsourcing and civilianising of military and security functions in war-torn Iraq.” (ZMA09VJA-PA)

„Therefore, [...] the convergence rate advantage outweighs this drawback.” (ZMA10OMT-ME)

„Finally, some of the participants were interviewed again a while after the recordings had been made.” (ZMA12OLA-LA)
5.1. Variation in cohesive devices: linkers’ functions

Based on 50 linguistic variables (conjuncts, adverbs) in 4 functional categories

<table>
<thead>
<tr>
<th>additive</th>
<th>adversative</th>
<th>causal</th>
<th>sequential</th>
</tr>
</thead>
<tbody>
<tr>
<td>also</td>
<td>nevertheless</td>
<td>because</td>
<td>firstly</td>
</tr>
<tr>
<td>moreover</td>
<td>although</td>
<td>therefore</td>
<td>secondly</td>
</tr>
<tr>
<td>furthermore</td>
<td>yet</td>
<td>consequently</td>
<td>previously</td>
</tr>
<tr>
<td>and</td>
<td>though</td>
<td>hence</td>
<td>afterwards</td>
</tr>
<tr>
<td>besides</td>
<td>but</td>
<td>then</td>
<td>eventually</td>
</tr>
<tr>
<td>actually</td>
<td>however</td>
<td>in this respect</td>
<td>finally</td>
</tr>
<tr>
<td>alternatively</td>
<td>actually</td>
<td>for this reason</td>
<td>lasty</td>
</tr>
<tr>
<td>regarding</td>
<td>in fact</td>
<td>on acc. of this</td>
<td>anyhow</td>
</tr>
<tr>
<td>similarly</td>
<td>instead</td>
<td>as a result</td>
<td>anyway</td>
</tr>
<tr>
<td>likewise</td>
<td>rather</td>
<td>on this basis</td>
<td>next</td>
</tr>
<tr>
<td>namely</td>
<td></td>
<td></td>
<td>at this point</td>
</tr>
<tr>
<td>in addition</td>
<td></td>
<td></td>
<td>to sum up</td>
</tr>
<tr>
<td>incidentally</td>
<td></td>
<td></td>
<td>in short</td>
</tr>
<tr>
<td>thus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for instance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in other words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>on t. other hand</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Counts:

- additive: also 5977, moreover 11, and 75499, besides 31, actually 594, alternatively 24, regarding 664, similarly 84, likewise 6, namely 508, in addition 86, incidentally 2, thus 959, for instance 73, in other words 47, on t. other hand 78
- adversative: nevertheless 55, although 480, yet 410, though 599, but 5243, however 933, actually 594, in fact 155, instead 292, rather 1004
- causal: because 2816, therefore 1513, consequently 87, hence 245, then 2369
- sequential: firstly 2816, secondly 1513, previously 87, afterwards 245, eventually 2369, finally 6, lastly 3

Counts are in ascending order from left to right.
## 5.1. Linkers by gender/language

relative/1M words and by 4 functional categories

<table>
<thead>
<tr>
<th></th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>additive</strong></td>
<td>28216</td>
<td>26788</td>
</tr>
<tr>
<td><strong>adversative</strong></td>
<td>3403</td>
<td>2956</td>
</tr>
<tr>
<td><strong>causal</strong></td>
<td>2901</td>
<td>1937</td>
</tr>
<tr>
<td><strong>sequential</strong></td>
<td>482</td>
<td>492</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Afrikaans</th>
<th>English</th>
<th>SAfrLang</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>additive</strong></td>
<td>28341</td>
<td>26787</td>
<td>29218</td>
</tr>
<tr>
<td><strong>adversative</strong></td>
<td>3234</td>
<td>3821</td>
<td>2704</td>
</tr>
<tr>
<td><strong>causal</strong></td>
<td>2208</td>
<td>2775</td>
<td>2484</td>
</tr>
<tr>
<td><strong>sequential</strong></td>
<td>564</td>
<td>520</td>
<td>314</td>
</tr>
</tbody>
</table>
5.1. Linkers by gender

relative/1M words and by 4 functional categories
5.1. Linkers by language

relative/1M words and by 4 functional categories
5.2. Linkers by discipline

relative/1M words and by 4 functional categories

<table>
<thead>
<tr>
<th></th>
<th>Anth</th>
<th>Chem</th>
<th>EngLit</th>
<th>Geo</th>
<th>Hist</th>
<th>Jour</th>
<th>Ling</th>
<th>MechE</th>
<th>CivE</th>
<th>Pol</th>
<th>Curr</th>
<th>Law</th>
<th>Psych</th>
</tr>
</thead>
<tbody>
<tr>
<td>additive</td>
<td>32788</td>
<td>19375</td>
<td>29210</td>
<td>28562</td>
<td>30021</td>
<td>26384</td>
<td>26455</td>
<td>21512</td>
<td>17945</td>
<td>30978</td>
<td>32107</td>
<td>27317</td>
<td>29390</td>
</tr>
<tr>
<td>adversative</td>
<td>2568</td>
<td>1352</td>
<td>6258</td>
<td>2392</td>
<td>3047</td>
<td>3035</td>
<td>4049</td>
<td>1383</td>
<td>1299</td>
<td>3485</td>
<td>2812</td>
<td>1617</td>
<td>3621</td>
</tr>
<tr>
<td>causal</td>
<td>1528</td>
<td>1188</td>
<td>4119</td>
<td>1311</td>
<td>2546</td>
<td>2256</td>
<td>2756</td>
<td>1579</td>
<td>1445</td>
<td>2647</td>
<td>2192</td>
<td>2132</td>
<td>3210</td>
</tr>
<tr>
<td>sequential</td>
<td>637</td>
<td>344</td>
<td>677</td>
<td>375</td>
<td>571</td>
<td>427</td>
<td>603</td>
<td>309</td>
<td>294</td>
<td>463</td>
<td>364</td>
<td>313</td>
<td>570</td>
</tr>
</tbody>
</table>
5.2. Linkers by discipline

relative/1M words and by 4 functional categories

- additive
- adversative
- causal
- sequential

Columns:
- Anth
- Chem
- EngLit
- Geo
- Hist
- Jour
- Ling
- MechE
- CivE
- Pol
- Curr
- Law
- Psych
6. Modals

6.1. Counting tool

```
C:\tmp\ZMA>analyze.pl dept/A wordlist_mod.csv
current file: ZMA06FA-AA.txt
current file: ZMA07NIW-AA.txt
current file: ZMA09VN-AX.txt
current file: ZMA09WHI-AA.txt
current file: ZMA10LFJ-AA.txt
current file: ZMA12OLE-AE.txt
can      385          0.141044756981715
could    192          0.0703392034817906
may      118          0.0432293021398505
might    46           0.016852100834179
must     59           0.0216146510699252
shall    2            0.000732700036268652
should   168          0.0615468030465667
will     333          0.121994556038731
would    333          0.121994556038731
Total number of words: 272963
C:\tmp\ZMA>
```
## 6.2. Modals by gender/language

<table>
<thead>
<tr>
<th>modal</th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td>can</td>
<td>1538</td>
<td>1738</td>
</tr>
<tr>
<td>could</td>
<td>825</td>
<td>585</td>
</tr>
<tr>
<td>may</td>
<td>739</td>
<td>556</td>
</tr>
<tr>
<td>might</td>
<td>317</td>
<td>191</td>
</tr>
<tr>
<td>must</td>
<td>496</td>
<td>424</td>
</tr>
<tr>
<td>shall</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>should</td>
<td>821</td>
<td>541</td>
</tr>
<tr>
<td>will</td>
<td>1362</td>
<td>1142</td>
</tr>
<tr>
<td>would</td>
<td>1209</td>
<td>849</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>language</th>
<th>Afrikaans</th>
<th>English</th>
<th>SAfrLang</th>
</tr>
</thead>
<tbody>
<tr>
<td>can</td>
<td>2020</td>
<td>1399</td>
<td>1127</td>
</tr>
<tr>
<td>could</td>
<td>695</td>
<td>888</td>
<td>637</td>
</tr>
<tr>
<td>may</td>
<td>726</td>
<td>499</td>
<td>707</td>
</tr>
<tr>
<td>might</td>
<td>288</td>
<td>311</td>
<td>157</td>
</tr>
<tr>
<td>must</td>
<td>527</td>
<td>325</td>
<td>569</td>
</tr>
<tr>
<td>shall</td>
<td>31</td>
<td>40</td>
<td>19</td>
</tr>
<tr>
<td>should</td>
<td>708</td>
<td>581</td>
<td>741</td>
</tr>
<tr>
<td>will</td>
<td>1587</td>
<td>1112</td>
<td>1052</td>
</tr>
<tr>
<td>would</td>
<td>878</td>
<td>1468</td>
<td>955</td>
</tr>
</tbody>
</table>
6.2. Modals by gender

The chart displays the frequency of modal verbs by gender. The modal verbs 'can', 'could', 'may', 'might', 'must', 'shall', 'should', 'will', and 'would' are shown, with data categorized by female and male. The vertical axis represents the number of occurrences, ranging from 0 to 2000. The chart illustrates gender differences in the use of these modal verbs in the text.
6.2. Modals by language
6.3. Modals by discipline

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>can</td>
<td>1410</td>
<td>1277</td>
<td>1702</td>
<td>1587</td>
<td>336</td>
<td>1915</td>
<td>1865</td>
<td>3163</td>
<td>2268</td>
<td>1193</td>
<td>1857</td>
<td>1511</td>
<td>1518</td>
</tr>
<tr>
<td>could</td>
<td>703</td>
<td>642</td>
<td>708</td>
<td>511</td>
<td>1267</td>
<td>714</td>
<td>998</td>
<td>726</td>
<td>129</td>
<td>552</td>
<td>553</td>
<td>890</td>
<td>1030</td>
</tr>
<tr>
<td>may</td>
<td>432</td>
<td>269</td>
<td>490</td>
<td>787</td>
<td>211</td>
<td>1017</td>
<td>1381</td>
<td>76</td>
<td>826</td>
<td>350</td>
<td>451</td>
<td>1254</td>
<td>1165</td>
</tr>
<tr>
<td>might</td>
<td>169</td>
<td>37</td>
<td>462</td>
<td>226</td>
<td>128</td>
<td>184</td>
<td>207</td>
<td>126</td>
<td>99</td>
<td>286</td>
<td>320</td>
<td>257</td>
<td>566</td>
</tr>
<tr>
<td>must</td>
<td>216</td>
<td>149</td>
<td>401</td>
<td>549</td>
<td>275</td>
<td>357</td>
<td>364</td>
<td>543</td>
<td>341</td>
<td>577</td>
<td>441</td>
<td>1796</td>
<td>208</td>
</tr>
<tr>
<td>shall</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>0</td>
<td>43</td>
<td>27</td>
<td>19</td>
<td>0</td>
<td>17</td>
<td>44</td>
<td>73</td>
<td>218</td>
<td>15</td>
</tr>
<tr>
<td>should</td>
<td>615</td>
<td>366</td>
<td>333</td>
<td>728</td>
<td>354</td>
<td>1158</td>
<td>1017</td>
<td>366</td>
<td>397</td>
<td>794</td>
<td>1387</td>
<td>1701</td>
<td>469</td>
</tr>
<tr>
<td>will</td>
<td>1220</td>
<td>605</td>
<td>1253</td>
<td>1145</td>
<td>626</td>
<td>1931</td>
<td>1871</td>
<td>739</td>
<td>1088</td>
<td>1385</td>
<td>1595</td>
<td>2015</td>
<td>1354</td>
</tr>
<tr>
<td>would</td>
<td>1220</td>
<td>560</td>
<td>1181</td>
<td>583</td>
<td>1670</td>
<td>952</td>
<td>1475</td>
<td>796</td>
<td>549</td>
<td>611</td>
<td>1450</td>
<td>1108</td>
<td>1581</td>
</tr>
</tbody>
</table>
6.3. Modals by discipline
6.3. Modals by discipline

![Bar chart showing modals by discipline]
7. Conclusion

case study:
- South Africa is unique in offering different cultural variations in one academic system, but Afrikaans, African, etc. style?

general:
- variations are culture-specific
- discipline culture > gender/language culture

metalanguage instruction on a functional basis?
- personal pronouns (explicitly categorical)
  > cohesion (overused through teaching?)
  > modality/hedging (complex, less awareness)

-- > more comparative research
References


English as an International Academic Language: Exploring conventions in different disciplines

English world-wide/Africa + corpus/computer applications + academic writing

Josef Schmied
English Language & Linguistics
Chemnitz University of Technology
http://www.tu-chemnitz.de/phil/english/ling/presentations_js.php
josef.schmied@phil.tu-chemnitz.de