THE PEDAGOGIC APPLICATIONS OF A LEXICAL DATABASE (SciE-Lex) TO SUPPORT RESEARCH WRITING

Natalia Judith Laso
Universitat de Barcelona

Suganthi John
University of Birmingham

Valladolid, 6 March 2015
OUTLINE

- Writing for research publication purposes. A reminder
- What is SciE-Lex?
- Developing SciE-Lex
- Validating SciE-Lex
- Aims of the study
- Pedagogic applications of SciE-Lex: Writing for publication Workshop
- Results from the Workshop
- Conclusions
- Limitations & Future research
Writing for Research Publication Purposes

So as to gain access to a discourse community and acceptance within it, one must become familiar with the “game strategies” (Etherington 2008) that govern scientific writing

a) better knowledge/command of the scientific discourse
b) improvement of one’s scientific production in English
c) THUS, better chance to publish one’s results internationally and gain recognition among community experts
The dominance of English in research settings

While there is a move to challenge the international academic community to support local publications in languages other than English, this contribution reports on how we hope to fill an immediate need for our Spanish medical researchers who have to publish in English in international journals for pragmatic reasons.
**SciE-Lex: A Lexicographic Tool for NNS Biomedical Writers**

- Lexical database of non-specialised (bio)medical terms ([http://www.ub.edu/grelic/eng/scielex2/scielex.html](http://www.ub.edu/grelic/eng/scielex2/scielex.html))

**AIMS:**

- provide the scientific community with *useful information* on the active use of *general terms* in the biomedical register
- help **NNS writers enhance their knowledge of collocations** in bio-medical English writing
Methodology: developing SciE-Lex

SciE-Lex

- Health Science Corpus compilation (HSC)
  - Impact scientific journals (biology, biochemistry, medicine)
  - 718 articles +4 million words

Information displayed:
- Lexicogrammatical and collocational information about the most common general terms, frequently used in the biomedical register
- List of bundles/phraseological units, prototypically used in the scientific discourse, their discourse function (“moves”) and textual distribution
METHODOLOGY: DEVELOPING SCIE-LEX (FIRST PHASE):

- *SciE-Lex* provides lexicogrammatical information about the most common collocations of general terms, frequently used in the biomedical register.

- Grammatical category (C)
- Inflected forms (M)
- Equivalent in Spanish (E)
- Morphosyntactic patterns of occurrences (C)
- Collocations (L)
- Examples of real use _extracted from the HSC_ (Ex)
- Usage notes (N)
SciE-Lex output (First phase)

http://www.ub.edu/grelic/eng/scielex2/scielex.html
**Methodology: developing SciE-Lex**

*SciE-Lex (Second phase)*

- Highlight **useful phrases and expressions** (bundles/phraseological units) used for **various rhetorical functions** in the scientific register

valuable resource for **NNS writers** to become aware of the **mechanics** that govern **academic writing**
METHODOLOGY: DEVELOPING SciE-Lex

Information displayed (Second phase):

- List of bundles/phraseological units, prototypically used in the scientific discourse
- Discourse function ("moves")
- Textual distribution
BUNDLES IN SCIE-LEX_OUTPUT
HTTP://WWW.UB.EDU/GRELIC/ENG/SCIELEX2/SCIELEX.HTML
**Discourse functions in SciE-Lex**


---

**Comparing and contrasting**

- to the same extent
- there was no difference
- no significant difference in
- no significant difference between
- When compared with
- as compared with
- in comparison with
- a comparison of
- was not significantly different
VALIDATION OF SciE-Lex

- Our evaluation consists of two stages:
  - Evaluation by experts (Pre-evaluation stage)
  - Evaluation by users
AIMS OF THE STUDY

• Explore if SciE-Lex, a lexical database of non-specialised biomedical terms, can be exploited by NNS writers to enhance their knowledge of collocations in bio-medical English writing

• Report on a writing workshop conducted for Spanish medical researchers at the University of Barcelona (April 2014) to support research writing
AIMS:

- to provide scientific writers the necessary skills to produce an academic research article using appropriate academic language and style
- to support writers with one article which we helped them bring up to publishable standard

PARTICIPANTS: NNS medical postgraduate students and researchers (n=10; collected data from 8)

Field of work: Life Sciences & Psychological Sciences
WRITING FOR PUBLICATION WORKSHOP

METHOD:

1. Participants sent workshop facilitators an 800 word draft of their writing.
2. Facilitators read and highlighted deviant (non-prototypical) collocations.
3. Facilitators devised a worksheet of nouns, verbs and adjectives and asked participants to search for their collocates in SciE-Lex.
4. Participants engaged in individual work on their writing based on data from the worksheet and the facilitators’ feedback.
5. Participants produced a second draft.
RESULTS FROM THE WORKSHOP

- Noun, verb and adjective collocations

<table>
<thead>
<tr>
<th>Nouns</th>
<th>advance, procedure, resistance, growth, study, finding, purpose, result, research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbs</td>
<td>appear, assess, consist, develop, seem</td>
</tr>
<tr>
<td>Adjectives</td>
<td>capable, responsible, related</td>
</tr>
</tbody>
</table>
# Results from the Workshop

<table>
<thead>
<tr>
<th>study (First draft)</th>
<th>study (Second draft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>For the present study, two isolates of Influenza A virus were used: an avian-origin LPAIV H5N2 subtype (A/Anas platyrhynchos/2420/2010) (H5N2) and a human-origin H1N1 subtype (A/Catalonia/63/2009) (pH1N1). (P1B)</td>
<td>In the present study, two isolates of Influenza A virus were used: an avian-origin LPAIV H5N2 subtype (A/Anas platyrhynchos/2420/2010) (H5N2) and a human-origin H1N1 subtype (A/Catalonia/63/2009) (pH1N1). (P1B)</td>
</tr>
</tbody>
</table>
We paid attention to two cell based binding affinity assays “MHC reconstitution assay” and “MHC-epitope stabilization assay”. (P2M)
## Results from the Workshop

<table>
<thead>
<tr>
<th>responsible (First draft)</th>
<th>responsible (Second draft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some of these outbreaks were responsible of avian-to-mammals transmissions, affecting also humans; thus, representing a threat to public health [2-4]. (P3J)</td>
<td></td>
</tr>
<tr>
<td>Some of these outbreaks were responsible for avian-to-mammals transmissions, affecting also humans; thus, representing a threat to public health [2-4]. (P3J)</td>
<td></td>
</tr>
</tbody>
</table>
# RESULTS FROM THE WORKSHOP

<table>
<thead>
<tr>
<th>capable (First draft)</th>
<th>capable (Second draft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Previously, our group identified the peptide VIN1, located in conserved regions of the influenza A virus hemagglutinin subunit 1, as <strong>capable to generate</strong> cross-reactive antibodies (abs) in pigs. (P3J)</td>
<td>1. Previously, we identified the peptide VIN1, located in conserved regions of the influenza A virus hemagglutinin subunit 1, as <strong>capable of generating</strong> cross-reactive antibodies (abs) in pigs. (P3J)</td>
</tr>
<tr>
<td>2. The re-introduction of genes <strong>capable to activate</strong> cell death in tumoral cells or genes that can modulate intrinsic cellular factors and eliminate cancer cells are among the most common approaches. (P4L)</td>
<td>2. The re-introduction of genes <strong>capable of activating</strong> cell death in tumoral cells or genes that can modulate intrinsic cellular factors and eliminate cancer cells are among the most common approaches. (P4L)</td>
</tr>
</tbody>
</table>
**RESULTS FROM THE WORKSHOP**

*consist (First draft)*

The last step **consists to complete** the staining and test sample on a Flow cytometer. (P2M)

*consist (Second draft)*

The last step **consists of completing** the staining and test sample on a Flow cytometer. (P2M)
CONCLUSIONS

• Positive results:
  – from the use of SciE-Lex: improved drafts (lexicogrammatical and phraseological issues)
  – from facilitators’ intervention: improved drafts at the textual level
  – from participants’ perception: a Satisfaction Questionnaire reveals that SciE-Lex provides these medical researchers with the language they need for their writing as well as a wider choice of language to improve their writing style

• So, SciE-Lex has proven to be successful in enhancing medical researchers’ knowledge of collocations in bio-medical English writing
LIMITATIONS AND FUTURE RESEARCH

• Limitations:
  – Participants depended on workshop facilitators to identify deviant collocations

• Future research:
  – Further similar workshops need to be conducted so that a broader community of medical researchers become familiar with SciE-Lex
  – Worksheets of exercises need to be developed and linked to SciE-Lex so as to help SciE-Lex users search this database for prototypical collocations
The support of the Spanish Ministerio de Ciencia e Innovación and FEDER is acknowledged (References HUM2007-64332/FILO and FFI2011-28947).
Thank you