

Feigning Formality

A CORPUS LINGUISTIC PILOT STUDY OF NOVICE ACADEMIC WRITERS

Robin Jungwirth, Nathaniel Woznicki, & Dr. Lynsey Wolter | Department of English—Linguistics



INTRODUCTION

Corpus linguistics makes use of large collections of texts to scientifically study the language used in texts and speech. Using corpus linguistics we are exploring the difference between novice and professional academic writers.

Most studies of academic writing have been based on professional writing. However, we are looking into novice academic writing. We hope to further explore how students learn to write in an academic register.

WRITING COURSES

Students at UWEC are placed into one of three required writing courses based on their determined level of writing. We took a look at:

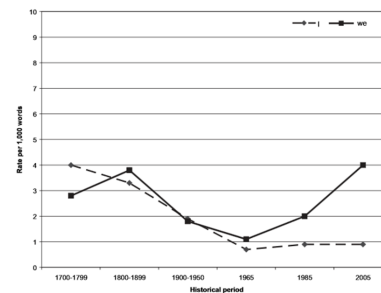
- ❖ WRIT 114 (bottom 10% based on placement test)
- ❖ WRIT 116 (middle 80% based on placement test)

PREVIOUS WRITING COURSE DATA

An anonymous post-evaluation of final papers from both writing courses revealed that students in WRIT 114 were considered as successful as students in WRIT 116.

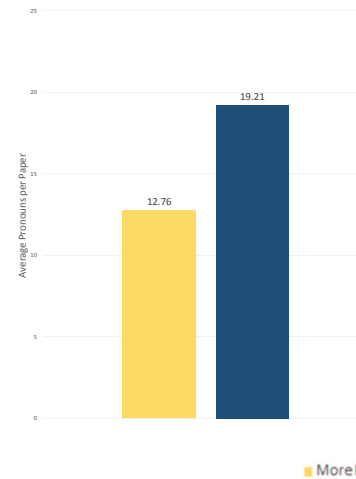
PREVIOUS RESEARCH

- ❖ Scholarly publications employ a compressed writing style
- ❖ There has been a possible rise in 1st-person pronoun usage since the 1960s
- ❖ Academic writing relies heavily on noun phrases to add information
- ❖ Passive voice usage is also more common in academic writing than in conversation

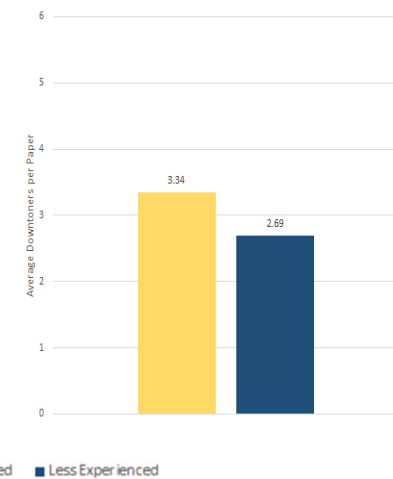


Historical Change of 1st Person Pronoun Usage in Academic Writing (Biber & Gray, 2010)

1st and 2nd Person Pronoun Usage



Downtoner Usage



SPECULATIONS AND FUTURE WORK

PRELIMINARY RESULTS

- ❖ Students in both courses had very low usage of hedges (*almost, maybe*) and downtoners (*barely, slightly*) in contrast to high usage based on findings by Biber & Gray (2010)
- ❖ Less-experienced students showed higher usage of 1st- and 2nd-person pronouns
- ❖ We considered, but did not find, notable evidence of differences in other factors such as word and text length

SPECULATIONS

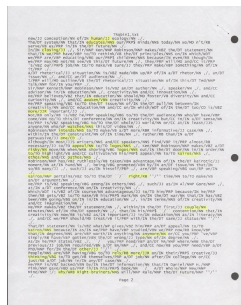
- ❖ Prescriptive transfer of writing expectations from secondary education, leading to certain trends in more-experienced writers that do not necessarily mirror professional academic writing
 - ❖ "Don't use *I* and *we*"
 - ❖ "Don't use words like *kind of*"
- ❖ Some level of assumed formality based on low frequency of hedges and downtoners

FUTURE WORK

- ❖ Investigate syntactically complicated features
 - ❖ Passive voice
 - ❖ Nominalizations
- ❖ Look further into expectations of novice writers from a secondary-education background
 - ❖ Sounding confident
 - ❖ Formality

REFERENCES

- ❖ Biber, D. & Gray, B. (2010). "Challenging stereotypes about academic writing: Complexity, elaboration, explicitness." *Journal of English for Academic Purposes*, 9, 2-20.
- ❖ Jarvis, S., Grant, L., Bikowski, D., & Ferris, D. (2003). "Exploring multiple profiles of highly rated learner compositions." *Journal of Second Language Writing*, 12, 377-403.
- ❖ Santorini, B. (1990). "Part-of-speech tagging guidelines for the Penn Treebank Project." Ms., University of Pennsylvania.



Sample of Tagger Accuracy Testing

WORKING WITH THE CORPUS

- ❖ Collected 200 anonymous rhetorical analysis papers (100 from each level) from E-Portfolio System
- ❖ Reformatted papers into plaintext documents
- ❖ Added part-of-speech tags using the Natural Language Toolkit (NLTK)
- ❖ Assessed accuracy through hand-correction of 7 Papers (93% accuracy on average) based on guidelines by Santorini (1990)
- ❖ Selected a number of linguistic features from Biber & Gray (2010) and Jarvis et al. (2003)
- ❖ Analyzed papers and procured preliminary word data using NLTK