



TECHNISCHE UNIVERSITÄT
CHEMNITZ

KDDE Modena , November 19, 2015

Researching Academia.edu: a Quantitative Corpus-Analysis of Modal Verbs and Personal Pronouns in Scientific Papers

Jessica Dheskali

1st Semester PhD Student

jessica.kuechler@s2009.tu-chemnitz.de



Introduction

Academia.edu:

- online platform for academics
- used by undergraduate students up to professors
- community that brings together people, interests, documents from the same specific area of research and knowledge
- open access- everyone can upload everything
- used to gain and track popularity in the academic community, share documents and get in touch with other academics



Methodology: General Information

Sample and Analysis

- corpus analysis with AntConc (2011)
- corpus of scientific articles from Academia.edu, mainly published (in journals)
- 4 affiliations: Faculty Member/ Department Member, Post-Doc, PhD Student, Graduate Student
- 7 different nationalities (British; German, Italian; Russian, Chinese; Ghanaian, Nigerian)
- 20 different Universities
- papers ranging from 2009-2015
- field: English Linguistics and Literature



Methodology: Corpus Compilation Problems

Text selection: platform problems

- different names for same university and department
- various departments with same or similar field
- unorganized and limited search options

Text provision: author problems

- paper written in mother tongue
- co-authors from different countries not mentioned
- paper not downloadable (Schmied 2013: 152)
- missing classification of document type



Methodology: Corpus Compilation Problems

Text processing: technical problems

- OCR programs needed for scanned papers (not all directly usable with AntConc)
- limitation in word count function of AntConc (2011)
- limitation in handling of annotated data in AntConc (2011)
- inevitable manual intervention in the process of transcription (McEnery, Ostler 2000: 411)

Corpus design issues:

- low representativeness of corpus (small sample) (Biber, Jones 2009: 1288)
- internally variation of characteristics of a text: differences in (argumentation) structure (Biber, Jones 2009: 1289)

English Language #3

MOSCOW STATE INSTITUTE OF INTERNATIONAL RELATIONS

4 People

11 Documents

Несобственно-прямая речь в английском языке: эволюция взглядов в западной лингвистике (1912-2012)

366

by Olga Blinova

Несобственно-прямая речь (НПР) – весьма распространенный стилистический прием в англоязычных текстах. Посвященные ему исследования появились около ста лет назад. За минувшее столетие взгляды лингвистов на НПР прошли через различные... [more](#)

Publication Date: Dec 2012

Publication Name: Блинова О.А. Несобственно-прямая речь в английском языке: эволюция взглядов в западной лингвистике (1912-2012). Вестник Новосибирского государственного университета. Серия: Лингвистика и межкультурная коммуникация. 2012. Т. 10. № 2. С. 93-102.

Research Interests: [Discourse Analysis](#), [Pragmatics](#), [Stylistics](#), [Linguistic stylistics](#), [Free Indirect Discourse](#), and 6 more

[View on elibrary.ru](#)

elibrary.ru/item.asp?id=18414279

[Download \(.pdf\)](#)

www.nsu.ru/xmlui/bitstream/handle/nsu/2373/11.pdf

[Bookmark](#)



The Notion of Free Indirect Discourse and its Use in Contemporary Journalism

988

by Olga Blinova

Free Indirect Discourse (FID) is traditionally considered a device limited to literary texts. This article argues that FID was widely used in journalist writing throughout the XXth century and is still popular today. The analysis focuses... [more](#)

Publication Date: Dec 2012

Publication Name: Blinova O.A. The Notion of Free Indirect Discourse and its Use in Contemporary Journalism. Humanities and Social Sciences Review ISSN: 2165-6258 :: Volume 01 :: Number 02 (2012) Copyright © 2012 by UniversityPublications.net

Research Interests: [Journalism](#), [Stylistics](#), [Discourse Stylistics](#), [Linguistic stylistics](#), [Ernest Hemingway](#), and 6 more

[View on universitypublications.net](#)

www.universitypublications.net/hssr/0102/html/HVD546.xml

[Download \(.pdf\)](#)

www.mgimo.ru/files2/2014_04/up11/file_084bc5f63d1c496fb2b9fa7297c4f5c0.pdf

[Bookmark](#)



О стилистическом потенциале несобственно-прямой речи в газетном тексте.

450

by Olga Blinova

Работа посвящена лингвистическим характеристикам и стилистическим функциям несобственно-прямой речи (НПР) в американской прессе. В статье затрагиваются как формальные характеристики НПР (дейксис, употребление времен,



ACADEMIA

Search...



HOME

ANALYTICS

SESSIONS

+ UPLOAD



Olga Blinova

Moscow state institute of international relations, English Language #3,... | Discourse Analysis +28

Address:

Moscow, Russia

62 Followers | 37 Following | 2,358 Total Views | top 5%

+ FOLLOW

MESSAGE

in 2

ALL 2 PAPERS IN ENGLI... 5 PAPERS IN RUSSI... 1 CONFERENCE PR... 1 BOOK REVIEWS 1 АВТОРЕФЕРАТ Д...

PAPERS IN ENGLISH



Speech and Thought Representation in Hemingway: The Case of Free Indirect Discourse

Bookmark | Download | 100 | More



The Notion of Free Indirect Discourse and its Use in Contemporary Journalism

Bookmark | Download | 988 | More

PAPERS IN RUSSIAN



Речевые характеристики персонажей-иностранцев в романе Э. Хемингуэя "По ком звонит колокол"



Methodology

Academia.edu-Corpus

- Total number of papers: 68
- Total number of words: 425 145

Nationality	Total nr. of words
Ghanaian	41274
Nigerian	50772
British	59033
Chinese	51885
German	76290
Italian	84035
Russian	61856

RQ1:

How is the distribution of personal pronouns among the different nationalities?

RQ2:

Which modal verbs are used to convey ideas, to put forward arguments and to make claims?



Results

	relative frequency per 100T words	
<i>may</i>	121,6	
<i>might</i>	40,5	
<i>should</i>	66,3	
<i>must</i>	50,6	
1st person SG	412	<i>I, mine, my, me</i>
1st person PL	387,2	<i>we, ours, our, us</i>
2nd person SG	145,6	<i>you, yours, your</i>

Table 1: Relative frequencies per 100,000 words
in the Academia.edu-Corpus



Results

	British	Chinese	German	Italian	Russian	Ghanaian	Nigerian	Average
<i>may</i>	118,6	100,2	98,3	204,7	87,3	164,8	69	121,6
<i>might</i>	59,3	52	72,1	35,7	25,9	12,1	7,9	40,5
<i>should</i>	40,7	86,7	69,5	53,6	74,4	63	84,7	66,3
<i>must</i>	40,7	63,6	23,6	31	27,5	75,1	130	50,6
1st person SG	833,4	354,6	458,8	643,8	646,7	697,8	283,6	412
1st person PL	340,5	235,1	347,4	363	679	438,5	291,5	387,2
2nd person SG	164,3	104,1	108,8	126,1	106,7	317,4	161,5	148,7

Table 2: Relative frequencies per 100,000 words in the Academia.edu-Corpus by the different nationalities



Results- Modal Verbs

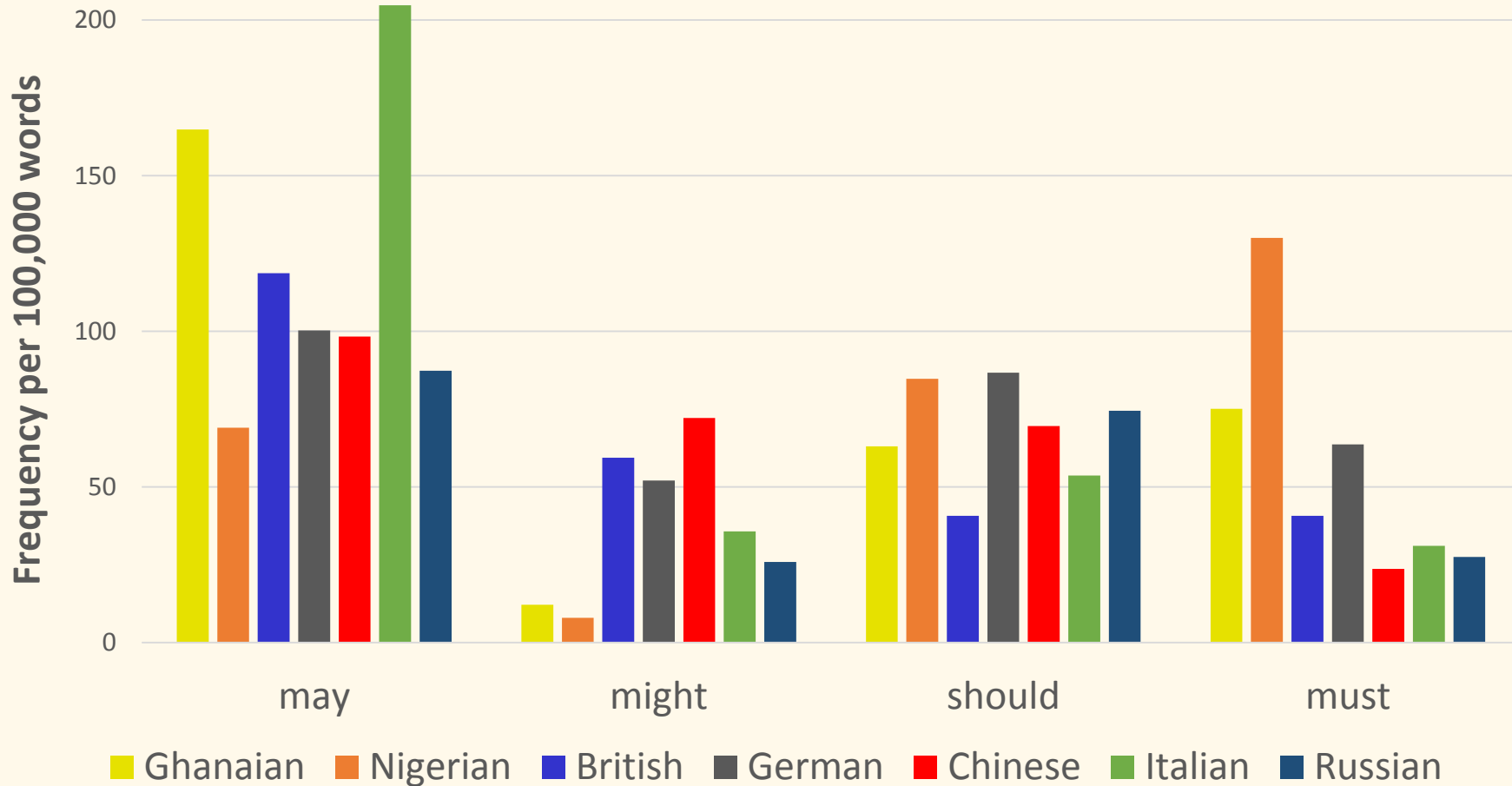


Fig. 2: Relative frequencies of modal verbs in the Academia.edu-Corpus



Results- Personal Pronouns

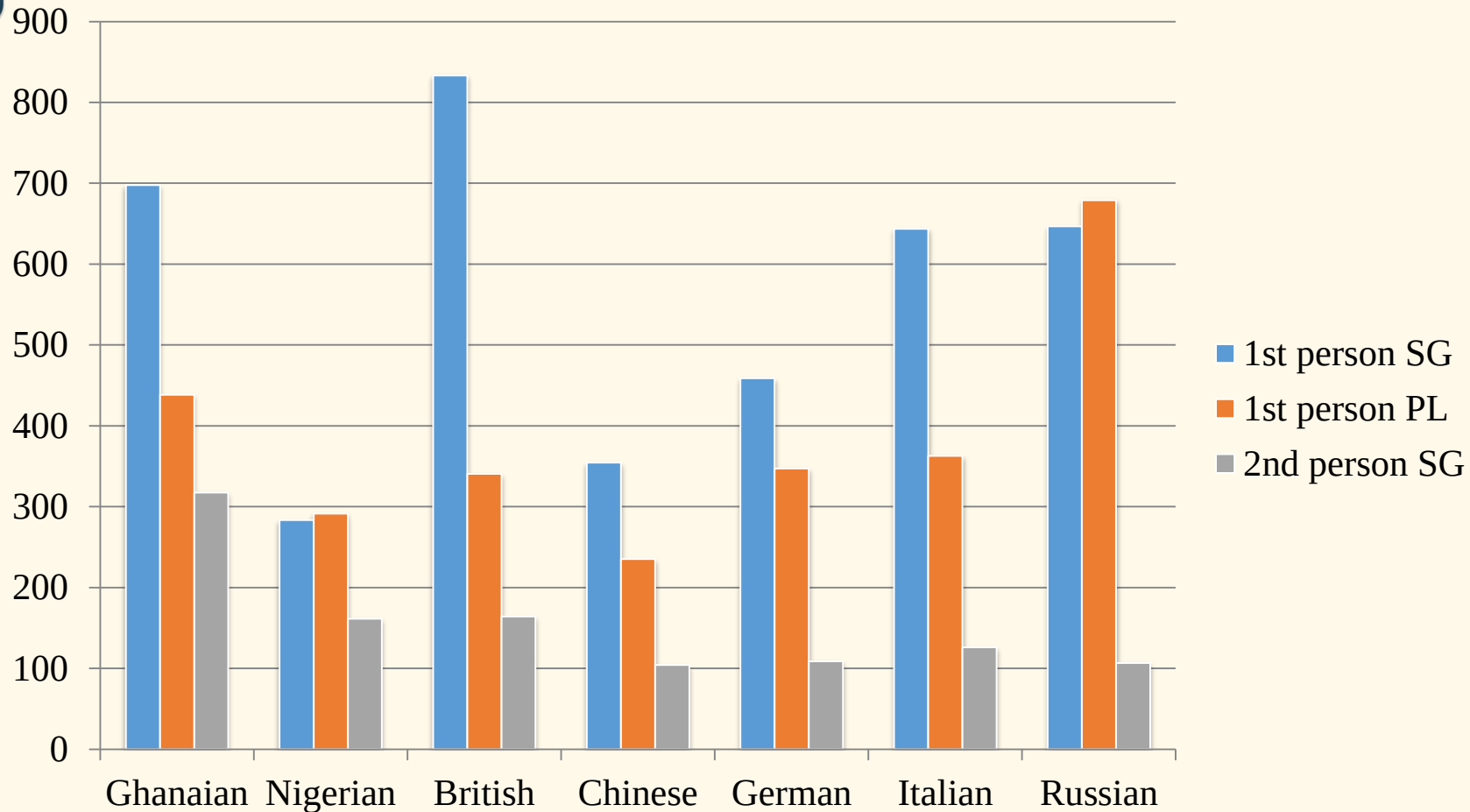


Fig. 2: Personal Pronouns in the Academia.edu-Corpus per 100,000 words



Discussion

- (1)What are the reasons for the high number of the 1st person PL pronoun in the Russian papers?
- (2)Is *we* rather used inclusively or exclusively in Russian papers? (Wales 1996: 63)
- (3)What are reason for highest number of hits of *must* in the Ghanaian and Nigerian papers?
- (4)How can we explain the high number of hits for *may* in Italian papers?



Results- without multi-authored papers (Russian)

	British	Chinese	German	Italian	Russian	Ghanian	Nigerian	Average
1st person PL	340,5	235,1	347,4	363	488,2	438,5	291,5	387,2

Table 2: Relative frequencies of 1st person PL pronoun per 100,000 words in the Academia.edu-Corpus the different nationalities

13 papers

- 6 with multiple authors (21029 words)
- 7 with single authors (40827 words)
- still highest number and above average



References

- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- McEnery, A.M. & Ostler, N. (200). A new agenda for corpus linguistics – Working With All of the World's Languages, *Literary and Linguistic Computing*, Volume 15, Issue 4, pp 401-418.
- Schmied, J. (2013). Academic Knowledge Presentation in MA theses: from Corpus Compilation to Case Studies of Disciplinary Conventions. *Brno Studies in English* 38, 149-164.
- Wales, K. (1996). *Personal Pronouns in Present-Day English*. Cambridge: Cambridge University Press.
- Biber, D., Jones, J.K. (2009). Quantitative Methods in Corpus Linguistics. In: Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics: An International Handbook. Vol.2*. Berlin/New York: Mouton de Gruyter, 987-1008.

Researching Academia.edu: a Quantitative Corpus-Analysis of Modal Verbs and Personal Pronouns in Scientific Papers

Jessica Dheskali

1st Semester PhD Student

jessica.kuechler@s2009.tu-chemnitz.de

List of Universities

University of Reading	Albert-Ludwigs-Universität Freiburg
Justus-Liebig-University Giessen	Università degli Studi di Torino
University of Pavia	Russian State University for the Humanities
Università Ca' Foscari Venezia	Saratov State University
Moscow State University	Shanghai Jiao Tong University
Moscow state institute of international relations	Hexi University
Peking University	University of Ibadan
Fudan University	SOAS University of London
University of Ghana, Accra, Legon	Queen Mary, University of London
University of Sussex	Russian Academy of Sciences